PROBING THE ROLES OF METAL BINDING LIGANDS IN CUPREDOXINS: INCORPORATING NONPROTEINOGENIC AMINO ACIDS INTO AZURIN AND CU_A AZURIN

BY

KEVIN M. CLARK

DISSERTATION

Submitted in partial fulfillment of the requirements
for the Doctor of Philosophy in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Yi Lu, Chair and Co-Director of Research
Professor Wilfred van der Donk, Co-Director of Research
Professor Antony Crofts
Professor Erik Jakobsson

ABSTRACT

Metalloproteins play important roles in biological systems since metal-binding sites are found in ~ 1/3 of structurally characterized proteins and in ~ 1/2 of all proteins. Despite progress, the role of many metal-binding sites remains poorly understood. Furthermore, their detailed geometric and electronic structure has yet to be fully elucidated. Site-directed mutagenesis using natural amino acids has continued to be the primary method of probing of metal-binding sites. However, it has become increasingly evident that site-directed mutagenesis, though powerful, cannot address the precise role of the metal ligands without changing multiple factors because of limited number of natural amino acids that allows isostructural substitutions. The relative paucity of functional group availability within the proteinogenic amino acids also limits ability to fine tune and expand protein's activity. To address the need to precisely determine specific metal binding ligand functionalities, and further efforts to engineer new function into proteins, expressed protein ligation (EPL) has been used to introduce non-proteinogenic amino acids to metalloproteins, such as the Type 1 (T1) copper site of azurin and the Cu_A site found in engineered Cu_A Azurin to aid the understanding of the roles of each ligand in this electron transfer protein.

This thesis covers the replacement of two copper-binding ligands in azurin, Cys112 and Met121, with nonproteinogenic amino acid analogues. Investigating the Cys112(S)-Cu Ligand-to-Metal Charge Transfer (LMCT) that gives azurin its intense blue color and its signature absorbance at 625 nm in the visible spectrum using two different analogues has been performed. Incorporating selenocysteine for Cys112 resulted in a new variant that maintained type 1 geometry , making it possible to elucidate the electron density of the site with more precision in the future. To further expand this concept, EPL was used to incorporate homocysteine (Hcy) to

investigate the rack hypothesis for the type 1 copper center. Insertion of a Cys derivative with an additional methylene group demonstrated that the type 1 site in azurin was capable of adjusting sufficiently through protein folding to maintain type 1 geometry and function. Together, these mutations provide insight into the unique relationship between the site's geometric features and its electronic features by minimally perturbing the site and selectively altering each feature individually.

Secondly, while conversion of blue copper sites with a weak axial ligand to green copper sites containing a medium strength axial ligand has been demonstrated in cupredoxins, converting blue copper sites to a red copper site with a strong axial ligand has not been reported. By replacing the weak Met121 thioether from azurin with a strong thiolate ligand from Hcy, a variant was created which had spectra (R_L= 1.5, and $A_{||}$ = 180 x 10^{-4} cm^{-1}) and a Cu-S(Cys) distance (2.22 Å) very similar to that of the red copper protein nitrosocyanin. The results strongly support the "coupled distortion" model, which helps explain axial ligand tuning of spectroscopic properties in cupredoxins, and demonstrate the power of using unnatural amino acids to address critical chemical biological questions.

Lastly, the unique Cu coordination of a backbone carbonyl in Cu_A azurin site was investigated using backbone carbonyl analogues of Glu. By replacing the carbonyl of the natural peptide linkage with an ester moiety, we have shown that the basicity of the coordinating carbonyl seems not to play a major role in either the spectroscopic or reduction potential features of the Cu_A site. These studies have provided a unique methodology for investigating metal ligand interactions beyond the standard side chain interactions.

In summary, a means to investigate the specific function of individual metal ligand residues in azurin and Cu_A azurin by using EPL to incorporate nonproteinogenic amino acid

analogues has been demonstrated. This work has resulted in new insights into the role of the metal ligands and has shaped future metalloprotein engineering efforts.

To my wife Eve whose support and love got me through this research.
I love you.

ACKNOWLEDGEMENTS

First I wish to thank my advisors, Wilfred van der Donk and Yi Lu. Thank you Wilfred for being a phenomenal advisor, mentor, and friend. Without your guidance, patience and willingness to help, I could not have completed the work described in this thesis. I will always be grateful for our conversations both about science and the nuances of life outside the lab that helped shape my journey through graduate school. Thank you Yi for being an exceptional mentor. Thank you for all of your guidance, understanding and teaching me how to be professional. Also, thanks to my committee members, Eric Jakobsson and Antony Crofts, who provided much support and guidance.

Thank you to all of the past and current members of the Lu group and van der Donk group whom I have had the honor working with for insightful conversations and helpful discussions. Thank you to Cyril Jacquot for being a wealth of knowledge and an immutable friend. Also, thanks to Noah Bindman, Patrick Knerr, and Lindsey Johnstone whose interesting chats always led to fun times in the lab. I will miss all the laughs and singing while in the lab. Thank you also to Lu group members Natasha Yeung, Yang Yu, Nicholas Marshall, Nathan Sieracki, and Kyle Miner for great discussions.

Thank you Martha Freeland, for being a lifeline during my tenure at UIUC. Thanks for being there for everything from supporting lab experiments to emotional support and medical aid. You will remain a great friend forever.

Thanks to my wife, Eve, and my dog, Einstein, who endured this long process with me, always providing laughs and love. Lastly, thanks to my family, especially my mother, Mary, and my extended family, my father-in-law, Albert, and mother-in-law, Tessie for all their support.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ACN.....acetonitrile
AcOH.....acetic acid
Asn.....asparagine
Asp.....aspartic Acid
ATP.....adenosine triphosphate
Az.....azurin
Boc.....tert-butyloxycarbonyl
CBD.....chitin binding domain
CcO.....Cytochrome c Oxidase
CV.....cyclic voltammetry
Cys.....cysteine
DBU.....1,8-diazabicyclo[5.4.0]undec-7-ene
DCM.....dichloromethane
DDAB.....dididecldimethylammonium bromide
DFT.....density functional theory
DIC.....N,N'-diisopropylcarbodiimide
DIPEA.....diisopropylethyl amine
DKP.....diketopiperazine
DMAP.....4-dimethylaminopyridine
DMF.....dimethylformamide
DNP.....dinitrophenyl
DTNP.....2,2'dithiobis(5-nitropyridine)
DTT.....dithiothreitol
EDCI.....1-ethyl-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDT.....ethanedithiol
EDTA.....ethylenediaminetetraacetic acid
ENDOR.....electron-nuclear double resonance
EPL.....expressed protein ligation
EPR.....electron paramagnetic resonance
ESEEM.....electron spin echo envelope modulation
ESI-MS.....electrospray ionization mass spectrometry
ET.....electron transfer
EtOAc.....ethyl acetate
EtOH.....ethanol
EXAFS.....extended x-ray absorption fine structure
FL.....full length
Fmoc.....fluorenylmethyloxycarbonyl
Gln.....glutamine
Glu.....glutamate
Gly.....glycine
HBTU.....O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate
HCTU.....O-(1H-6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate
Hcy.....homocysteine

HEPES ..... 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His ..... histidine
HOBt ..... 1-hydroxybenzotriazole hydrate
IMPACT ..... intein mediated purification with an affinity chitin-binding tag
LB ..... luria broth
Leu ..... leucine
LMCT ..... ligand to metal charge transfer
Lys ..... lysine
MALDI-MS ..... matrix-assisted laser desorption ionization mass spectrometry
MeCN ..... acetonitrile
MES ..... 2-(N-morpholino)ethanesulfonic acid
MESNa ..... sodium 2-mercaptoethanethiosulfonate
Met ..... methionine
MOPS ..... 3-(N-morpholino)propanesulfonic acid
MPAA ..... mercaptophenyl acetic acid
mRNA ..... messenger RNA
MWCO ..... molecular weight cutoff
NCL ..... native chemical ligation
NEB ..... New England Biolabs
NMA ..... N-methylmercaptoacetamide
NMM ..... N-methylmorpholine
NMR ..... nuclear magnetic resonance
NOR ..... nitrous oxide reductase
NpsCl ..... 2-nitrophenyl sulfenyl chloride
PAM 4-hydroxymethyl-phenylacetamidomethyl
PEG ..... polyethylene glycol
PGE ..... pyrolytic graphite edge
PMB ..... *p*-methoxybenzyl
PNBP ..... 4-(4-nitrobenzyl) pyridine
PMSF ..... phenylmethanesulfonylfluoride
PS ..... polystyrene
RNA ..... ribonucleic acid
RP-HPLC ..... reverse phase high performance liquid chromatography
rR ..... resonance raman
RT ..... room temperature
SDM ..... site directed mutagenesis
SECIS ..... selenocysteine insertion sequence
Sec ..... selenocysteine
Ser ..... serine
SPPS ..... solid phase peptide synthesis
TCEP ..... tris-(2-carboxyethyl)phosphine
TFA ..... trifluoroacetic acid
TFE ..... trifluoroethane
TFMSA ..... trifluoromethanesulfonic acid
THF ..... tetrahydrofuran
Thr ..... threonine

TIPS ..... triisopropylsilane
TLC ..... thin layer chromatography
TMSOTf ..... trimethylsilyl trifluoromethanesulfonate
TOF ..... time of flight
tricine ..... N-(tri(hydroxymethyl)methyl)glycine
GdmCl ..... guanidinium hydrochloride
TRIS ..... tris(hydroxymethyl)aminomethane
tRNA ..... transfer RNA
Trp ..... tryptophan
VMD ..... Visual Molecular Dynamic
WANG ..... *p*-benzyloxybenzyl alcohol resin
WT ..... wild type
XAN ..... xanthenyl
XANES ..... x-ray absorption near edge structure
XAS ..... x-ray absorption spectroscopy
Z ..... benzyloxycarbonyl

LIST OF FIGURES

LIST OF SCHEMES

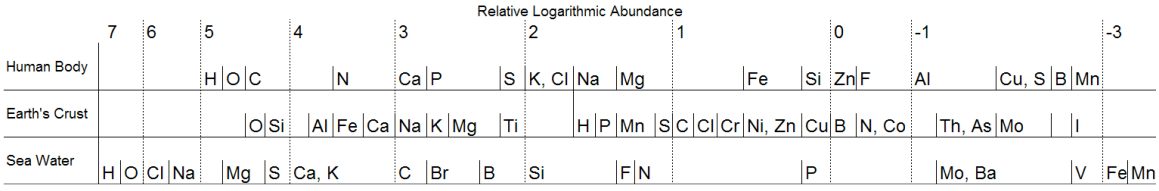
LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 TRANSITION METALS IN BIOLOGY

The distribution of transition metals found in the human body is very different compared to the relative abundance of metals found in nature (Figure 1.1).[1,2] The distribution of metal ions found in the human body more closely resembles that of seawater than of earth.[3,4] Moreover, of all the metals found in nature, only nine (iron, zinc, copper, molybdenum, vanadium, manganese, chromium, nickel, and cobalt) are biologically relevant.[5] These biologically relevant metals, when used enzymatically, are bound by a large series of proteins containing unique ligand sets creating functions without which life would not exist. These proteins, termed metalloproteins, constitute greater than half of all the proteins found in life.[5] With functions ranging from oxygen transport to electron transfer and redox chemistry, metalloproteins are intimately involved in a host of cellular processes absolutely necessary for life.



Element	Concentration in Sea Water (M x 10-8)	Concentration in Human Plasma (M x 10-8)
Fe	0.005-2	2230
Zn	8	1720
Cu	1	1650
Ni	0.5	4.4

Figure 1.1: Comparison of metal distribution in the human body compared to the earth and seawater. A) Relative logarithmic abundance. B) Relative concentrations of ions found in seawater compared to humans.

Proteins with standard natural functional groups and no metal cofactors are limited when performing reactions requiring electron equivalents. Metalloproteins however, possess the distinct advantage of efficiently shuttling or storing electrons needed for chemical processes by using metal cofactors. Without metalloproteins' capability to shuttle and store reducing and oxidizing equivalents, key cellular processes such as the generation of ATP would be dismantled.

How do metalloproteins manage to create such a diverse set of functions with such a limited supply of metals and metal ligands? One method to expand metalloprotein functionality is utilization of identical ligand sets with slightly different geometries in different protein scaffolds. Having a different scaffold with the same ligand set retains the overall protein function while the minor changes in geometry and bond length of the site can allow fine-tuning to accommodate new physiological function. Additionally, substitution of metal ligands with other natural amino acids with different functional groups can expand metalloprotein function, imparting new protein function.

Arguably, nature uses different metals to accomplish identical functions. For example, three different metal containing families (heme iron, non-heme iron, and copper proteins) are shown in Table 1.1. This small sample set demonstrates that using considerably different metals and metal ligand sets retains similar chemistry. Using this rationale, it is clear that cells can adapt to conditions of their environment by substituting available metals for metals that are limited to maintain necessary cellular functions.

	Heme Fe	Non-heme Fe	Cu
Dioxygen transport $DM^n + O_2 \rightarrow M^{(n+1)}{}_D(O_2{}^D)$	Hemoglobin Myoglobin	Hemerythin Myohemerythin	Hemocyanin
Monooxygenases $X + O_2 + 2e^- + 2H^+ \rightarrow XO + H_2O$	Cytochrome P-450	Soluble MMO	Tyrosinase
Dioxygenases $X + O_2 \rightarrow XO_2$	indoleamine 2,3-dioxygenase tryptophan 2,3-dioxygenase prostagladin H synthase	lipoxygenase α-ketoglutarate dependent dioxygenases	Quercetinase
Oxidases $O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$ $XH_2 + O_2 \rightarrow X + H_2O_2$	Cytochrome c Oxidase	Ribonucleotide Reductase Cofactor generation	Cytochrome c Oxidase Laccase
Peroxidases and Catalases $H_2O_2 + XH_2 \rightarrow X + 2H_2O$	Catalase Horseradish Peroxidase Cytochrome c peroxidase		
Superoxide Dismutase $2O_2^- + 2H_2^+ \rightarrow O_2 + H_2O_2$		Fe-SOD	Cu-SOD
Electron Transfer	Cytochromes c	Ferredoxins ([Fe-S] proteins)	Cupredoxins (Type 1 and Cu_A)

Table 1.1: Functional diversity in Fe and Cu proteins.

1.2 COPPER PROTEINS

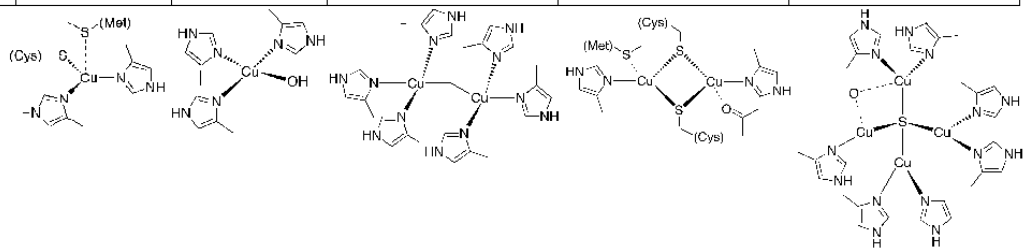
Copper is one of nine biologically relevant metals and a trace metal essential for all living organisms.[6] Copper metalloproteins function in oxidation, and dioxygen transport and participate in fast electron transfer. These different functions are achieved by modifying the Cu ligand set in their respective protein scaffolds (Table 1.1). The diverse range of activities of Cu proteins resulted in development of a classification scheme dividing Cu metalloproteins into six major divisions based primarily on the number of Cu ions bound by the protein (Table 1.2).[7] Each of the divisions is further subdivided by their biological function. The mononuclear copper proteins function in electron transfer and catalysis. The binuclear copper proteins function in electron transfer

and O_2 activation and transport (Table 1.2). Trinuclear and tetranuclear copper proteins are involved in catalytic reactions and are newer members and therefore less characterized.

While copper proteins involved in reactions other than electron transfer are interesting and have been reviewed many times,[8-14] this thesis focuses solely on the copper sites involved in electron transfer; the type 1 copper or blue copper proteins and the Cu_A or purple copper proteins. The low Cu^+ to Cu^{2+} redox potential (~150 mV) combined with the abundance of Cu found in human serum (Figure 1.1b) results in a series of proteins with high Cu binding affinity and extremely fast electron transfer rates.

Table 1.2: Copper protein classification based on active site. The basic structure representation for each copper site is shown below the table.

	Mononuclear		Dinuclear		Tetranuclear
Biological function	Electron transport	Catalysis	O_2 transport and activation	Electron transport	Catalysis
Examples	Azurin Plastocyanin Laccase Nitrite reductase	Amine Oxidase Galactose oxidase Superoxide dismutase Nitrite reductase	Hemocyanin Tyrosinase Laccase Catechol oxidase	Cyt c oxidase Nitrous-oxide reductase Menaquinone NO-reductase	Nitrous-oxide reductase
UV-vis spectroscopy	Strong absorption 600 nm	Weak absorption 700 nm, Forbidden d-d transistion at 600 nm	Intense absorption at 300 nm, Weak absorption at 700 nm	Strong absorption at 480 nm and 530 nm	Strong absorption at 650 nm
EPR spectrum	4-line, small hyperfine	4-line, large hyperfine	No EPR signal	7-line, small hyperfine	2x4-line, small hyperfine
Active site geometry	Distorted tetrahedral	Distorted tetragonal	Tetragonal	Trigonal planar	m_4-S^{2-} tetracopper cluster



1.3 TYPE 1 COPPER PROTEINS

General Protein and Metal Binding Site Structure

Copper proteins that function solely in electron transfer are known as cupredoxins. Cupredoxins are Cu proteins that contain the type 1 copper site or the Cu_A site. Type 1 copper proteins contain a uniquely conserved Greek-key β-barrel fold (eight β-strands that form two β-sheets folded around a center α-helix) (Figure 1.2 A). The mononuclear Cu binding site is located within the C-terminal β-strands. The mononuclear site binds Cu with three conserved ligands in a trigonal plane.[15] The trigonal plane consists of two equatorial histidine residues and a single equatorial cysteine. The Cu ion is also coordinated by at least one additional axial ligand (Figure 1.2 B) and these proteins share a conserved short Cys(S)-Cu(II) bond (2.1-2.5 Å) resulting in strong covalent character and a very intense absorbance at ~600 nm.

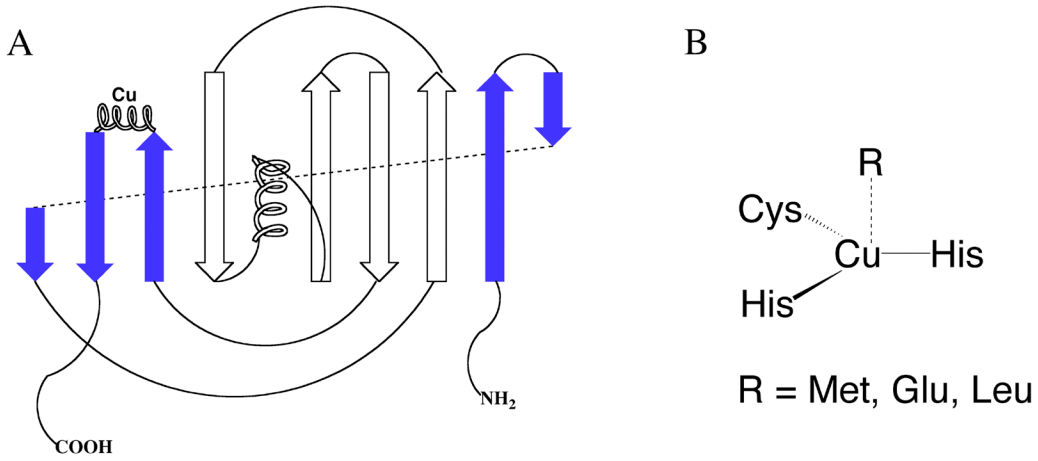


Figure 1.2: The general type 1 Cu protein fold and metal binding site. A) The Greek-key β-barrel fold of cupredoxins. B) The general type 1 Cu site. See text for details.

Though all type 1 copper proteins contain identical equatorial ligand sets, the amino acid in the axial ligand position differs. The axial ligand currently is considered to play a significant role in the magnitude of the redox potential of each type 1 copper protein.[16,17] Though the axial ligands typically only result in minor structural differences, they can significantly alter the electrostatics of the site. Because of these differences, the type 1 copper proteins have been subdivided into three classes; Class 1, Class 2, and Class 3 (Figure 1.3), based on the identity of the axial ligand present. The class 1 proteins, such as amicyanin, have a methionine residue in the axial position. Class 2 proteins, such as stellacyanin, have a non-methionine residue such as glutamine (stellacyanin) or leucine (laccase) in the axial position. Class 3 proteins, while also containing a methionine in one axial position contain an additional axial interaction with a backbone carbonyl. Azurin, the prototypical Class 3 type 1 copper protein, has a fifth coordination with the backbone carbonyl of Gly45. The importance of the differing axial position functional groups will be addressed further in Chapter 4.

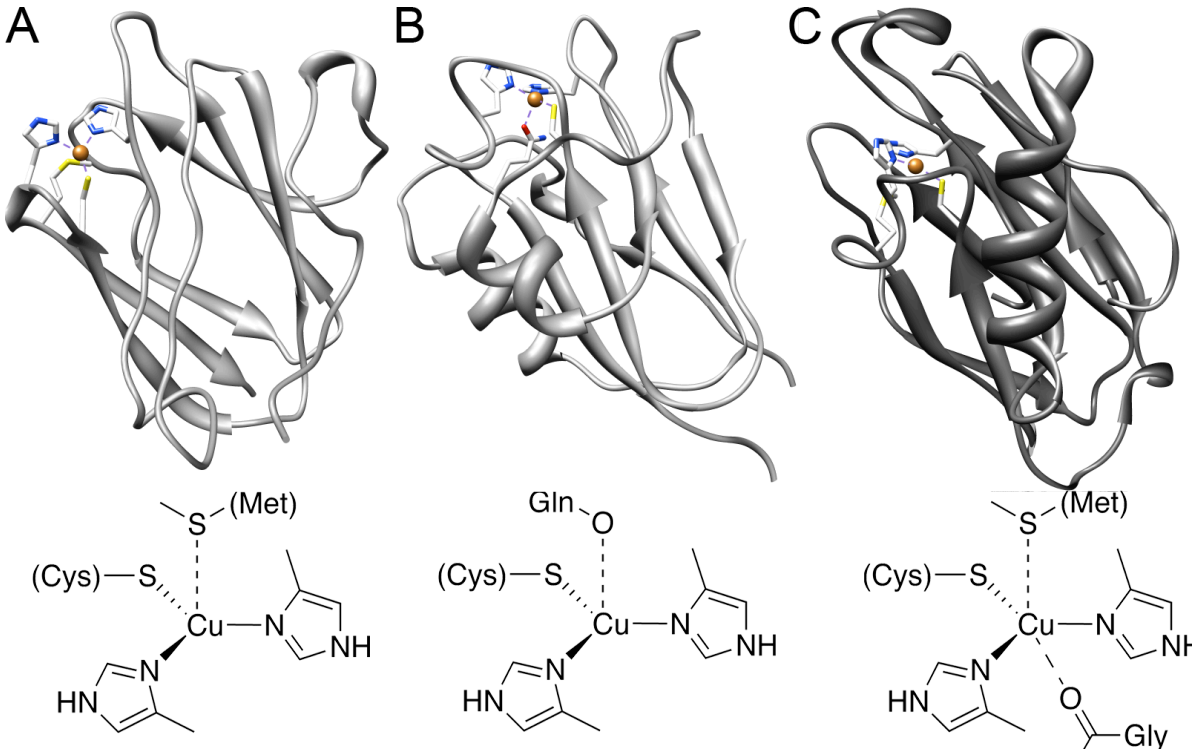


Figure 1.3: Type 1 Cu protein classes - A) Class 1 - amicyanin (PDB Code: 1AAC). The metal site has a Met as the axial ligand. B) Class 2 - stellacyanin (PDB Code: 1JER). The metal site has a Gln as the axial ligand. C) Class 3 - azurin (PDB Code: 4AZU). The metal site has a Met in one axial position as well as a Gly backbone carbonyl as a secondary axial ligand.

1.4 AZURIN FROM *PSEUDOMONAS AERUGINOSA*

Pseudomonas aeruginosa azurin is a well studied, prototypical type 1 Cu protein scaffold that has advanced the understanding of cupredoxin function. Azurin is a class 3 blue copper protein and therefore has five ligands to coordinate the Cu ion. Within the conserved trigonal plane lie two histidine residues (His 46, His117) and a single cysteine (Cys112). Methionine (Met121) is one of the axial ligands and serves as a weak axial ligand. Opposite to Met121 is a second weak axial interaction with the backbone carbonyl of Gly45 (Figure 1.4).[7,18,19]

The type 1 copper proteins contain an unusually short Cys(S)-Cu(II) bond distance of approximately 2 Å. The Cys(S)-Cu(II) bond distance in azurin is 2.1 Å and

confers strong covalent character giving an intense ligand-to-metal charge transfer (LMCT) band stemming from the Sγ(Cys112) 3pπ → Cu(II) d_x^2-d_y^2 charge transfer (CT) transition.[7] The LMCT imparts azurin with its blue color and strong absorption at 625 nm (Figure 1.5 A). The LMCT is also responsible for the unique EPR spectrum seen for azurin (Figure 1.5 B). Characterized by 4 lines of equal intensity with narrow parallel hyperfine splittings ($A_{\|}$ = 60 Hz), the EPR spectrum differs greatly from "normal" copper coordination compounds (Figure 1.5 B). Typical synthetic copper compounds have much larger hyperfine splittings ($A_{\|}$ > 160 Hz) and do not always mimic the type 1 copper site features well (Figure 1.5).[7]

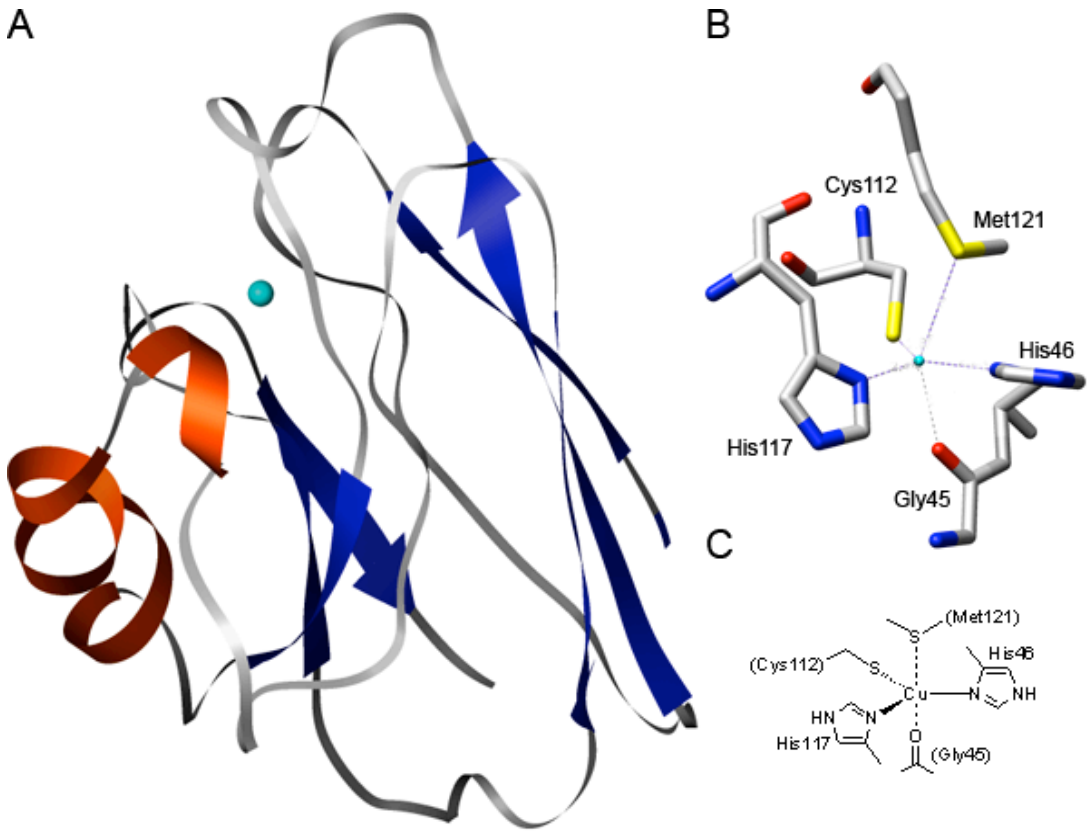


Figure 1.4: *Pseudomonas aeruginosa* azurin. A) The full-length holo-azurin protein (PDB Code: 4AZU). B) Type 1 Cu site of azurin. C) Schematic diagram of the type 1 site found in azurin.

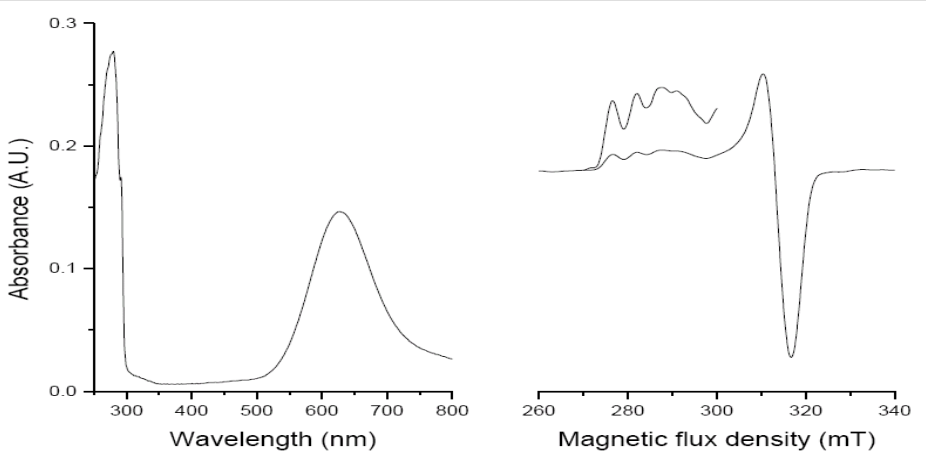


Figure 1.5: *Pseudomonas aeruginosa* azurin spectroscopic characterization. A) UV-vis spectroscopy shows the LMCT band at 625 nm. B) X-band EPR shows narrow parallel hyperfine splittings

The Role of Cys112 in Azurin

The equatorial Cys112 and the axial Met121 ligands in azurin are suggested to be responsible for the large increase in redox potential and more efficient ET properties over "normal" copper complexes. The type 1 copper site in azurin has been exhaustively modified using site-directed mutagenesis and its variants have been extensively studied using a broad range of spectroscopic[20-30] and crystallographic techniques.[31-35] These thorough investigations have shown that Cys112 is essential for protein function.

Early studies established the loss of the azurin spectral properties when Cys112 was substituted with other natural amino acids differing greatly in size and electronics. In contrast, substitution of the remaining three ligands, His46, His117, and Met121 primarily yields proteins dominated by the strong LMCT and blue color.[22,36-39] However, the presence of the Cys112 thiol is not a requisite for copper binding. A single Cys112 mutant, Cys112Asp, retained the ability to bind copper,[40] though all the spectral features including the distinct LMCT band at 625 nm disappeared, and the ET properties of azurin were greatly reduced.[40,41] Additional spectroscopic studies established the original type 1

copper site had been transformed into a type 2 "normal" copper site with the Cu bound by a bidentate Asp112.[40,42] Thus while maintaining the ability to bind Cu within the mononuclear site, the protein was no longer capable of performing typical azurin functions.

The Cys(S)-Cu(II) bond length in azurin is exceptionally short at ~2.1 Å. At such a short distance, it can be expected that the bond between the Cys thiolate and the Cu(II) metal would be a covalent interaction. Resonance Raman (rR) and X-ray absorption spectroscopy (XAS) were very useful in confirming the covalent nature of the S-Cu bond.[43-47] The high Cu-S stretching frequency in rR combined with the XAS data and DFT calculations has verified the S-Cu as highly covalent and has shown the amount of S character (a direct measure of covalent interaction) to be ~38%.[48,49] Though these studies have provided support for a covalent interaction between Cys112 and Cu, the relative contributions of electron density of the metal-ligands and the metal remain debated. By analyzing the two protons attached to Cβ of Cys112 using ENDOR spectroscopy some light has been shed onto this predicament.[50] Using ENDOR, the spin density of the sulfur-centered orbitals was determined to be approximately 30-40%.[48-50] The remaining electron density resides on the Cu and and is spread weakly over the other ligands in the trigonal plane.

The Role of Methionine 121 in Azurin

Methionine 121 was thought to be the ligand responsible for tuning the redox behavior of azurin and has undergone extensive study by site-directed mutagenesis and substitution with nonproteinogenic amino acids.[16,17,36,51,52] Initial crystal structures of

azurin placed the Met121-Cu bond at a distance greater than 3 Å, suggesting little interaction of the axial ligand with the Cu ion.[33,53,54] However, more recent structures and computational calculations placed the bond length closer to ~2.5 Å, supporting the idea that the Met121 thioether can interact with the Cu.[23,55,56] An extensive list of mutants was created to better probe the axial ligand interaction with Cu.[26,39,52,57-59] While the majority of these mutants display very similar spectroscopic signatures as WT azurin, a tremendous range of reduction potentials was observed with azurin Met121 variants and has substantiated the role of Met121 in determining the redox behavior of azurin.

The Met121 variants provided evidence for control of the redox couple in azurin and have provided evidence to support that the axial ligand may also play multiple other roles in azurin. One of the proposed roles of Met121 is protecting the site from exogenous small molecule interactions. Some azurin Met121 mutations have transformed the azurin type 1 site into mimics of similar type 1 copper proteins such as nitrite reductase or laccase,[21,60-64] supporting the idea that the axial ligand may play a role in protecting the copper ion from exogenous ligands and water.[65-67] Such development adds an additional method to modulate the redox behavior of azurin and its type 1 copper cousins.[29,51] Moreover, it has been proposed that the axial ligand may function in maintaining correct active site geometry.[68-71]

1.5 Cu_A Azurin

General Protein and Metal Site Structure

The Cu_A site is found in terminal oxidases such as cytochrome c oxidase (CcO) and nitrous oxide reductase and is intricately involved in electron transfer and the proton

shuttle for the generation of ATP. While the type 1 copper proteins have high structural homology, proteins containing the Cu_A copper site do not share a common protein fold. The Cu_A site geometry however is well conserved with only minor axial ligand changes.

The Cu_A site is currently the only electron transfer site known to contain binuclear Cu and represents the only known Cu-Cu bond within a protein. It displays entirely different geometry than its electron transfer cousins, the type 1 copper proteins.[72] The Cu_A site displays geometry more similar to dinuclear non-heme iron proteins and Fe-only/FeNi hydrogenases (though they do not share common ligand sets) than to other Cu electron transfer sites. The 2 Cu ions sit in a rigid diamond core consisting of μ2 bridging thiolates from two cysteine residues that hold the Cu ions in a diamond shape (Figure 1.6, C/D). The diamond core helps limit rearrangement during the redox event and contributes to exceptionally fast electron transfer kinetics and abnormally small reorganization energy.[54,73-77] The small reorganization energy is imparted by the site's mixed valence of Cu(1.5) • Cu(1.5). The oxidation to the Cu(II) • Cu(II) state is not favorable, thus limiting the range of reduction potentials allowed by the site.[73-75] The wide range of reduction potentials found in type 1 Cu proteins are not found in Cu_A proteins since the Cu_A site function is more strictly conserved than type 1 sites.

Two additional equatorial histidine residues complete the geometries of the Cu_A site diamond core. Similar to the type 1 copper site, each copper ion also has additional axial ligands such as methionine thioether or backbone carbonyl interactions from nearby residues (Figure 1.6). These equatorial ligands have been shown not to be the major determinant of the redox potential of Cu_A sites.[78]

The study of the Cu_A site in its native scaffold is challenging since native proteins containing Cu_A sites are either membrane bound or very large (e.g. bovine CcO is >200 kDa). Moreover, these proteins contain multiple other metal-binding sites giving rise to cross signals that interfere with data collection and make interpretation difficult. Thus typical spectroscopic techniques that are used to characterize metalloprotein active sites cannot be used in many Cu_A scaffolds. While model structures using synthetic chemistry were successful in providing insight into the structure of the Cu_A site,[79,80] the structures were unsuccessful in providing insight into Cu_A site functionality because study inside a protein scaffold is needed to fully understand the site's delicate nature.

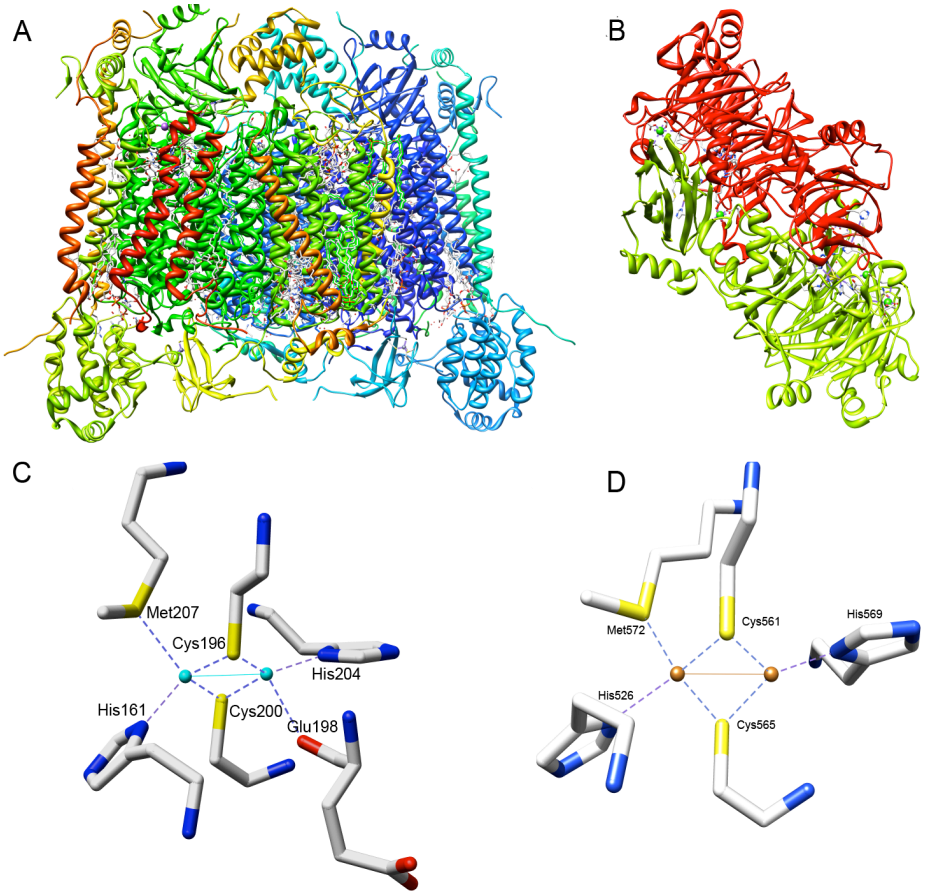


Figure 1.6: Structural comparison of Cytochrome c Oxidase and nitrous oxide reductase. A) Crystal structure of bovine heart C*c*O (PDB ID: 2DYR).[81] B) Structure of nitrous oxide reductase from *P. denitrificans* (PDB ID: 1FWX).[82] C) Cu site in CcO. D) CuA Site in nitrous oxide reductase.

The Cu_A Azurin Protein Model

Using protein scaffolds other than native structure is a powerful tool to understand metalloprotein active site dynamics. Since most Cu_A scaffolds have unwanted, interfering cross signals from other metal sites, development of a new scaffold that retained only the Cu_A signal would be extremely useful for characterization of the Cu_A site. Though the scaffolds of type I and Cu_A proteins differ dramatically, the sites reside in similar cupredoxin folds. Thus a new model system was proposed that would eliminate the unwanted cross signals of other metal sites in the native scaffold while maintaining native Cu_A signals.[81] However, finding a model system that mimics not only the active site but also maintains the native spectroscopic signatures is difficult. The primary sequences of the azurin type 1 Cu binding loop and the Cu_A Cu binding loop found in *P. denitrificans* C*c*O were compared and had sequence similarities.[83]

```
WT Azurin  AECSVDIQGN DQMQFNTNAI TVDKSCKQFT VNLSHPGNLP KNVMGHNWVL 50
WT CuA Az  AECSVDIQGN DQMQFNTNAI TVDKSCKQFT VNLSHPGNLP KNVMGHNWVL 50

WT Azurin  STAADMQGVV TDGMASGLDK DYLKPDDSRV IAHTKLIGSG EKDSVTFDVS 100
WT CuA Az  STAADMQGVV TDGMASGLDK DYLKPDDSRV IAHTKLIGSG EKDSVTFDVS 100

WT Azurin  KLKEGEQYMF FCT-FPG-HS ALMKGTLTLK 128
WT CuA Az  KLKEGEQYMF FCSELCGINH ALMKGTLTLK 130
```

Figure 1.7: Sequence overlay of azurin and Cu_A azurin. The red box highlights the differences in the active site sequence.

Engineering a Cu_A site into the azurin scaffold required extensive mutation. Creating a Cu_A site in azurin was accomplished by substituting residues in the type 1 copper binding loop in wild type (WT) *P. aeruginosa* azurin with the residues in the Cu_A binding loop of *P. denitrificans* C*c*O (Figure 1.7).[83] This new methodology, termed loop-mutagenesis, created a new, engineered protein (Cu_A Az) (Figure 1.8).[83-85] The native Cys112 Cu ligand was not replaced to maintain continuity within the azurin scaffold

while attempting to limit protein misfolding. To maintain the conserved CxxxC motif found in many Cu_A proteins a second Cys residue was introduced at position 116.

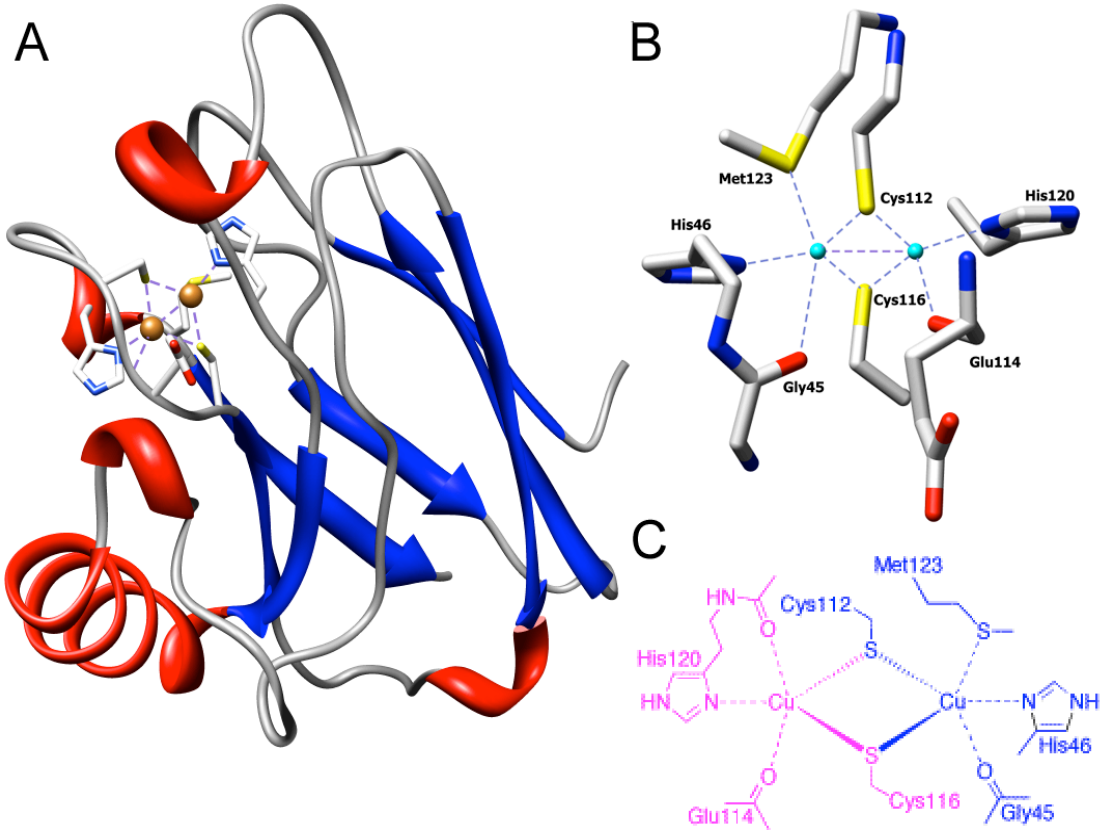


Figure 1.8: Engineered Cu_A azurin. A) Full length Cu_A Az (PDB: 1AAC). B) The di-nuclear Cu_A Az site. Residues are numbered in bold. C) Schematic representation of the Cu_A Az site. The residues highlighted in blue represent the ligands found naturally in the type 1 copper site in azurin. The residues in purple were engineered to fulfill the Cu_A site.

Spectroscopic characterization of Cu_A Az revealed the engineered protein would provide an excellent model for the Cu_A site.[85] The LMCT bands at 480 nm and 520 nm with a d-d transition at 750 nm are well in line with spectra seen in other Cu_A sites (Figure 1.9 A). Furthermore, the characteristic 7 line EPR spectra of Cu_A sites were also seen in the engineered scaffold (Figure 1.9 B).[85] While allowing for simpler spectral characterization of the Cu_A site in an easy to express scaffold, new difficulties in

characterizing the Cu_A Az chimera were encountered. Specifically new pH dependencies not found in native Cu_A scaffolds resulting in some type I copper signal were observed. However, these complications are seen at pH extremes and before Cu binding is complete. Inspection of the engineered site shows how some type 1 characteristics could remain when the engineered Cu_A site is overlaid with the original type 1 site in azurin (Figure 1.8 C). If the site were imaginarily divided, one of the bound Cu ions would retain most of the type 1 ligand set (Figure 1.8C, blue) and the second bound Cu would contain a new ligand set differing from azurin (Figure 1.8C, purple). Thus when titrating with Cu, it is conceivable that the original type I site may momentarily form before the second reduced Cu ion binds, temporarily displaying type I characteristics.

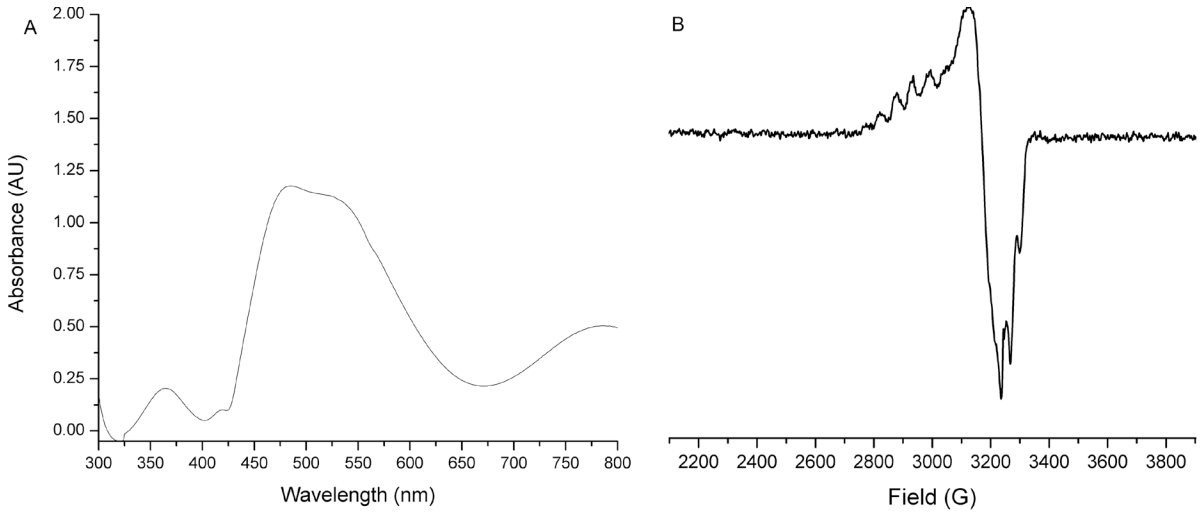


Figure 1.9: Spectroscopic Characterization of Cu_A azurin. A) UV-vis spectrum. B) X-band EPR spectrum.

A study that substituted the bridging Cys residues in Cu_A Az with Ser revealed a separation of spectral features, neither of which gave the engineered Cu_A signals.[86] Rather, substitution of Cys116 resulted in a protein with type 2 characteristics and a

yellow color while substitution of Cys112 resulted in a protein with type 1 characteristics and blue color. This separation of spectral features is not found in Cys-to-Ser mutations of the natural *P. denitricans* C*c*O site. Therefore, while the engineered protein acts as a good mimic of a WT Cu_A site, the Cu_A Az scaffold is not without its own limitations of study.

1.6 Methods of Nonproteinogenic Amino Acid Incorporation

While site-directed mutagenesis provided substantial information towards understanding both the type 1 Cu and Cu_A electron transfer sites, studying the role of each residue of these sites is difficult since the paucity of functional groups in the 20 naturally occurring amino acids (of which only seven are available for metal binding) severely limits the coordination chemistry. The selective incorporation of nonproteinogenic amino acids that contain functional groups not available in proteinogenic amino acids into the type 1 Cu and CuA sites could fine tune metalloprotein function. Moreover, the roles of the individual Cu metal binding ligands could be better elucidated by substitution with isosteric nonproteinogenic amino acid analogues. Incorporation of nonproteinogenic amino acids, however, cannot be done with standard mRNA translation.

Natural mRNA translation uses a specific three-nucleotide codon for the recruitment of aminoacylated tRNA factors for correct proteinogenic amino acid incorporation. However, these codons are specific for their cognate tRNA synthetases, eliminating the ability to use the current genetic code to incorporate nonproteinogenic amino acids. Many tRNA synthetases are so intimately designed to accept their cognate amino acid that they reject similar natural isosteric analogues through a molecular

"double sieve" mechanism.[87,88] These specificity filters limit misacylation of tRNAs and create another barrier to using cellular machinery for the incorporation of nonproteinogenic analogues. Though difficult and much more involved than standard proteinogenic amino acid incorporation, natural biosynthetic methodologies exist for incorporation of nonproteinogenic amino acids.

1.7 Biosynthetic Routes for Nonproteinogenic Amino Acid Incorporation

Translational Recoding

The twenty known proteinogenic amino acids were long thought to be the only set of building blocks used for protein biosynthesis. However, the discovery that selenium was essential for the function of two proteins, glycine reductase and glutathione peroxidase, began a search for the function of Se in a cellular environment, leading to the discovery of selenocysteine (Sec, U), the 21st amino acid.[89,90] The revelation that Se was directly incorporated into these proteins as the novel amino acid, selenocysteine, was a profound discovery.[91] Other recently discovered naturally incorporated novel amino acids have questioned the limitation of the genetic code and have shown that amino acids outside the standard twenty proteinogenic amino acids are readily available in some species.[92-94]

To date, the incorporation of extra proteinogenic amino acids has relied on the ability to recode known stop codons, a technique known as translational recoding or codon suppression. The incorporation of Sec is currently the best-understood route and though the route uses available cellular machinery, the incorporation of Sec is far from simple. Sec incorporation is done through a translational recoding event at a UGA opal

stop codon and requires four gene products, *selA*, *selB*, *selC*, and *selD*.[95] Improper read-through of the stop codon results in truncated variants of the desired translated protein.

A series of *cis* and *trans* factors are required to correctly read-through the UGA codon for Sec insertion. Some parallels in how Sec and the standard 20 amino acids are incorporated do exist. Both methods have dedicated tRNAs (*selC* for Sec) and EF-Tu-like factors (*selB* for Sec) for ribosome recruitment. However, key differences exist. The Sec-tRNA$^{[Ser]Sec}$ used for translational incorporation is different from other standard tRNAs in two ways. First, the tRNA^{Sec} acceptor arm is longer than traditional proteinogenic tRNAs. Secondly, Sec-tRNA^{Sec} is also unlike the other bacterial and eukaryotic tRNAs that are acylated with the correct amino acid and require no further chemistry. Sec-tRNA^{Sec} is charged with a serine residue (Ser-tRNA^{Sec}) that is converted through a series of steps to selenocysteine using the gene products of *selA* (a selenocysteine synthase) and *selD* (a selenophosphate synthetase).[96-98] While this type of acylation chemistry to synthesize the desired amino acid on tRNA occurs in certain methanogens it is not a common methodology for tRNA acylation.[99]

As shown in Figure 1.10, Seryl-tRNA synthetase initiates the process for the incorporation of Sec by charging tRNA^{Sec} with Ser. Once acylated, the synthesis of Sec from Ser begins through the generation of selenomonophosphate (SeP_i) by selenophosphate synthetase (SelD).[95-98,100] Once generated, selenocysteine synthase (SelA) uses the SeP_i to begin converting the Ser residue to selenocysteine.[101,102] When the ribosome reaches the UGA codon used for Sec incorporation, a downstream RNA hairpin termed the SECIS element (Selenocysteine Insertion Sequence) signals for the recruitment of a *trans* acting protein, SelB required for incorporation of Sec. SelB is an

EF-Tu homolog containing a SECIS binding loop region in its C-terminus. SelB interacts with the SECIS element and the ribosome causing the ribosome to momentarily pause on the UGA stop codon.

Once paused on the UGA codon, the ribosome's interaction with SelB causes it to refrain from initiating a translational abort. This pause also provides time for the EF-Tu homologous domain of SelB to recruit the correctly acylated Sec-tRNA^{Sec} where it then undergoes GTP hydrolysis, incorporating Sec at the UGA codon.[100,101] The efficiency and success of Sec incorporation is highly dependent on the availability of all of the required *cis* (SECIS) and *trans* (*selA/B/C/D*) elements. Missing a single component greatly decreases incorporation efficiency.

While selective recombinant expression of selenoproteins was accomplished, expression is limited due to the translation recoding needed for UGA and differences in SECIS element recognition. Differences in both the location of and the distance between the SECIS loop and the UGA codon within prokaryotic and eukaryotic systems greatly complicate expression system development.

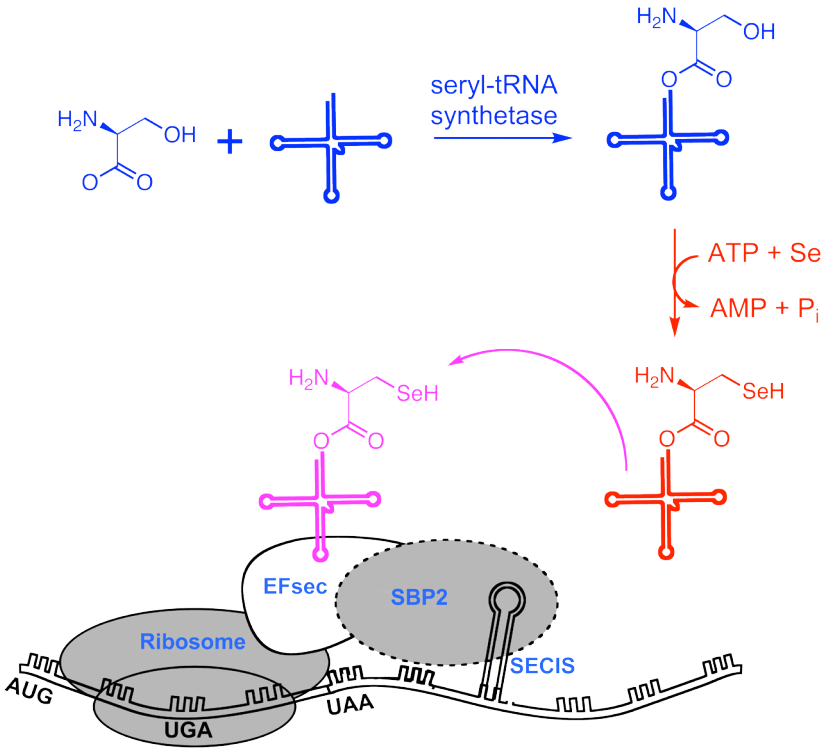


Figure 1.10: General Scheme for incorporation of selenocysteine using SECIS elements. Amino-acylation of tRNA with serine (shown in blue) is the first step of synthesis followed by the installation of the selenol shown in red. A specific tRNA is used for the recruitment to the ribosome. See text for more details.

In prokaryotic systems, the SECIS loop is located in the coding region at a distance of ~20 nucleotides from the UGA codon. The amino acids following the UGA codon are encoded and translated into the complete protein product.[102] In contrast, the eukaryotic and archea SECIS loops have variable distances from the UGA codon, are found exclusively in the 3`-untranslated region of the gene, and are consequently not translated into the protein product.[103] This difference in proximity of the SECIS elements has limited the ability to translationally encode selenoproteins in *E. coli* recombinant expression systems. However, Hilvert and coworkers have recently used a redesigned SECIS element to recombinantly incorporate Sec into cytochrome P450cam to investigate heme axial ligand interaction with Cys.[104] This report currently represents the

only fully engineered SECIS element for expression of a eukaryotic selenoprotein product in a recombinant expression system.

Another potential limitation using SECIS elements is the inability to incorporate multiple Sec residues per protein. Incorporation of multiple Sec residues in an engineered protein sequence may be required to gain function or to study the specific roles of residues in the metal ligand set.

Nonsense Suppression

Nonsense suppression techniques are a comparable technique to natural Sec incorporation. Similar to translational recoding of the opal stop codon for Sec incorporation, some *E. coli* strains are known to naturally recode the amber stop codon (UAG). The amber stop codon is the least used stop codon by *E. coli.* Recoding the UAG stop codon provides an efficient methodology for expressing proteins containing specific nonproteinogenic amino acid substitutions.[105,106]

The incorporation of nonproteinogenic amino acids using nonsense suppression requires many aspects to be effective.[107] A unique codon must be available or created that can only be recognized by an exogenously expressed aminoacyl-tRNA synthetase, ensuring correct and efficient translation and subsequent incorporation of a nonproteinogenic amino acid. However, to be efficient at incorporation, the engineered synthetase cannot acylate endogenous tRNAs. Therefore, for a nonsense suppression system to work, an exogenously expressed, non-host synthetase/tRNA pair must efficiently interact with the host ribosome while not interfering with the host's natural translational machinery. Discovering an aminoacyl-tRNA synthetase/tRNA pair that

suites the above requirements requires a tremendous amount of selection.[105] Adding to the difficulty, the desired nonproteinogenic amino acid must easily cross the membrane to be incorporated if it cannot be made *in vivo*, and the analogue must not serve as a substrate for *any* endogenous host synthetases.[105-106]

Though a complicated and difficult challenge to overcome, Schultz and coworkers have developed a toolkit to produce orthogonal-tRNA:aminoacyl-tRNA-synthetase pairs capable of incorporating a broad range of nonproteinogenic amino acids.[107-111] By exposing the tRNA/tRNA-synthetase pairs from *Methanococcus jannashii* to rounds of positive and negative selection, the successful incorporation of nonproteinogenic analogues was achieved site-specifically into a range of recombinantly expressed proteins in *E. coli*. While greater than 30 different types of probes have been incorporated into proteins using nonsense suppression techniques, the successful generation of a tRNA/tRNA-synthetase pair remains a limitation of the technique.

Auxotrophic Substitution

Using auxotropic bacteria is an older method for incorporating nonproteinogenic amino acid analogues. An auxotrophic bacterium is an engineered strain that has lost the ability to produce a particular amino acid, requiring external addition of the amino acid for incorporation into *any* growing polypeptide chain. The incorporation of seleno-amino acid analogues such as selenomethionine using *E. coli* auxotrophs has long been performed to solve phasing problems in protein crystallography. This methodology has also been applied to the use of cysteine auxotrophic bacteria to substitute cysteine residues with selenocysteine for similar purposes.[112,113] However, using bacterial

auxotrophs for proteinogenic amino acid replacement with nonproteinogenic amino acids is not without its limitations.

Using auxotrophs for replacement of amino acids with their isosteric analogues suffers from two key limitations. First, the method replaces all proteinogenic amino acids found in the protein with the analogue and therefore typically alters more than one single site of interest. Secondly, though deficient in the synthesis of a particular amino acid, cells do find alternative routes to make amino acids and recycle those already in the cell. Consequently, complete controlled replacement with analogues using this technique has yet to be accomplished, with typical replacement percentages between 60-80%.[114,115] Incomplete replacement results in multiple different protein isoforms, complicating characterization and ultimately reducing the technique's usefulness to directly probe the individual functionalities of residues.

1.8 Synthetic and Semi-Synthetic Routes for Nonproteinogenic Amino Acid Incorporation

While biosynthetic routes for the incorporation of nonproteinogenic amino acids represent one method for direct substitution of amino acid residues, a chemical approach can alternatively create a regimented methodology that can be repeated for a multitude of different nonproteinogenic amino acids, with potentially limitless possibilities for selective location of incorporation.

Standard Solid Phase Peptide Synthesis

Using Solid Phase Peptide Synthesis (SPPS) to directly synthesize full-length proteins containing nonproteinogenic amino acids was one of the first chemical methodologies applied to studying protein function with nonproteinogenic amino acids. Using PEG or PS resins containing different linkers, peptides and full-length proteins of interest can be synthesized step-by-step using a purely chemical methodology. While seeming to be the perfect means for the incorporation of nonproteinogenic amino acids, since virtually any functional group can be incorporated into the synthesized peptide chain, SPPS suffers from major limitations.

The number of residues that are consecutively coupled to the peptide chain limits solid phase peptide synthesis. To achieve high yields each coupling step of the synthesis requires high efficiency (>99%). However, as the peptide chain elongates, secondary structure can form and the peptide can fold upon itself making additional couplings to the free amine more difficult. This ultimately decreases the coupling efficiency, lowering the final yield. Over the span of a long peptide sequence, these relatively small decreases in coupling efficiency result in decreased peptide yields and increased peptide impurities. Peptide sequences made by SPPS therefore typically do not exceed 70 amino acids in length, limiting the technique's usefulness in the synthesis of full-length proteins, as very few protein sequence are less than 100 amino acids in length.

Native Chemical Ligation

Native Chemical Ligation (NCL) was developed to use the advantages of SPPS but enable synthesis of longer peptides.[116] NCL uses thioester-mediated ligation of two

peptide sequences to make longer peptides that are inaccessible by standard SPPS techniques. Two peptides of interest, in which one or both contain a single or multiple nonproteinogenic amino acids, are synthesized separately by standard SPPS and purified to homogeneity. The proposed mechanism of NCL is shown in Figure 1.11. NCL begins when the free N-terminal cysteine thiol of *Peptide 2* attacks the C-terminal thioester of *Peptide 1* via a chemoselective transthioesterification (Figure 1.11, Step 1) forming a transient thioester intermediate (Figure 1.11). The intermediate undergoes a spontaneous S-N acyl rearrangement (Figure 1.11, Step 2), forming the native peptide bond with a Cys at the linkage site.

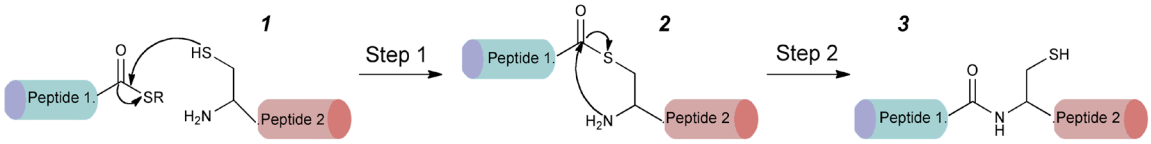


Figure 1.11: Native chemical ligation. In step 1, a transthioesterification of the two peptides yields the intermediate product ***2***. A spontaneous N-S acyl rearrangement yields the final product ***3*** containing a peptide bond.

NCL can provide peptide sequences of significantly longer length than those available by standard SPPS. Preparation of synthetic peptide-alkyl esters to facilitate transthioesterification has been achieved by alkylation of peptide thiocarboxylates and by direct synthesis.[116-118] Though many of these thioester derivatives are stable, their reactivity at pH 7 is extremely low resulting in long reaction times and increased side-reactions. To promote faster transthioesterification, external thiol catalysts such as benzyl mercaptan, thiophenol, and 2-mercaptoethanesulfonate (MES) have been added.[119,120] Though these external catalysts can decrease reaction times, they suffer in the presence of

hindered thioesters. Recent work by Kent and coworkers examined the proposed mechanism of NCL by investigating a wide range of thiol catalysts to attain faster reaction kinetics and minimize side-reactions commonly seen in NCL.[121] It was subsequently determined that aryl-thiols capable of rapid thiol-exchange yet retaining high reactivity with Cys for transthioesterification, provided the highest protein yields.

The requirement of a Cys residue at the junction site is a potential limitation of NCL. To perform NCL the native protein sequence must contain an easily accessible Cys residue to carry out the ligation or a Cys residue must be engineered into the sequence at a position to make a suitable length peptide by SPPS. Techniques exist to reduce Cys residues to Ala using Raney nickel reduction.[122] Moreover, different chemistries such as using auxiliaries to facilitate ligation rather than a Cys residue have been developed to avoid the obligate need for a Cys residues in NCL by.[123]

Expressed Protein Ligation

Expressed Protein Ligation (EPL) is a recombinant extension to NCL.[116,118,124,125] In NCL, synthetic installation of a C-terminal thioester moiety is required to facilitate the transthioesterification needed to afford the full-length (FL) polypeptide. As this step can be relatively difficult to accomplish, suffers from low yields, and is limited by the length of peptide that can be synthesized using SPPS, an alternative methodology to create a C-terminal thioester moiety would be advantageous. As an alternative to NCL, EPL uses a modified version of protein splicing to generate protein-thioesters (Figure 1.12).

RNA splicing is a natural process used to increase the number of gene products from a single gene sequence. Similarly, protein splicing is a natural process used to

increase the number of protein products from a single mRNA sequence. Protein splicing gives rise to two protein products, a spliced variant usually an enzymatic protein and an intein that can be a homing endonuclease to limit protein translation. Protein splicing is similar to RNA gene splicing. However, protein splicing is found in prokaryotes, eukaryotes, and archea whereas gene splicing is not found in prokaryotes. The comparison between the splicing methods is depicted in Figure 1.12. Genes in eukaryotes are divided into two separate elements, introns and exons and by removing the introns and linking the exons, different mRNA variants are translated leading to different protein products. Likewise, mRNA that will be transcribed for protein splicing contains two separate elements, inteins and exteins. The inteins catalyze their excision from the protein sequence and link the exteins together to create fully functional proteins.

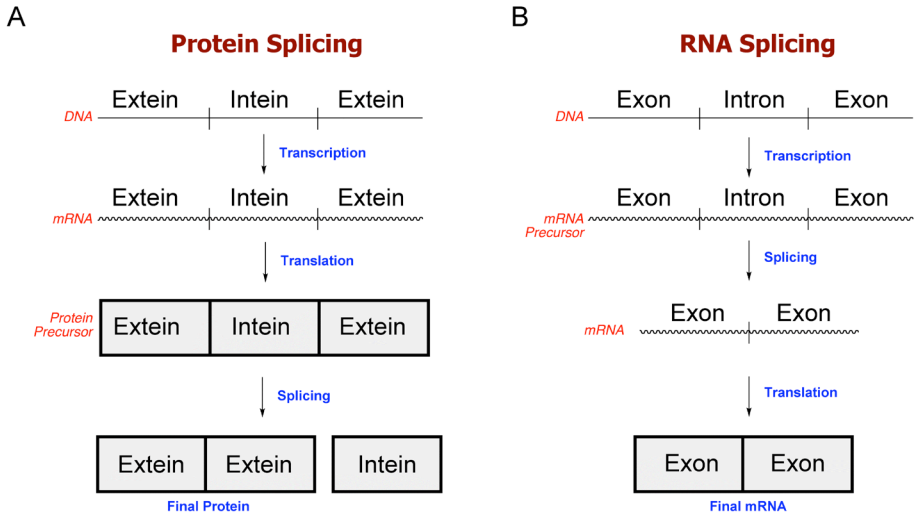


Figure 1.12: Comparison of protein and RNA splicing pathways. A) Protein splicing follows standard DNA transcription and mRNA translation to create a precursor protein that is then spliced using a self-excising intein to create the final protein. B) RNA splicing follows standard DNA transcription to make a mRNA precursor from which the exons are spliced, the introns removed, and standard mRNA translation then results in the final protein.

Following translation the exteins are joined by inteins that catalyze their *ipso* excision.[126-129] The splicing event at the intein/extein junction is catalyzed using internal Cys or Ser residues (Figure 1.13). Protein splicing begins when the intein undergoes a N-S/O acyl rearrangement, creating a transient ester at the C-terminal of the *first extein* (Figure 1.12, Step 1).[128,129] Attack of this ester *via* a transesterification by a Cys or Ser residue at the N-terminal of the *second extein* causes the splicing of the adjacent exteins by an intermediate ester linkage, but the intein remains attached to the N-terminus of extein 2 (Figure 1.12, Step 2.)

Cyclization of an Asn residue *via* attack of the amide by the amine side chain mediates excision of the intein. The intein-maleimide is then released from the ligated exteins (Figure 1.13, Step 3). The final FL polypeptide is afforded following a spontaneous S/O-N rearrangement, to create the amide bond between the adjacent exteins. Typically the extein is a protein needed for enzymatic reactions whereas the intein-malemide is hydrolyzed to the acid and used as an endonuclease to regulate protein expression.

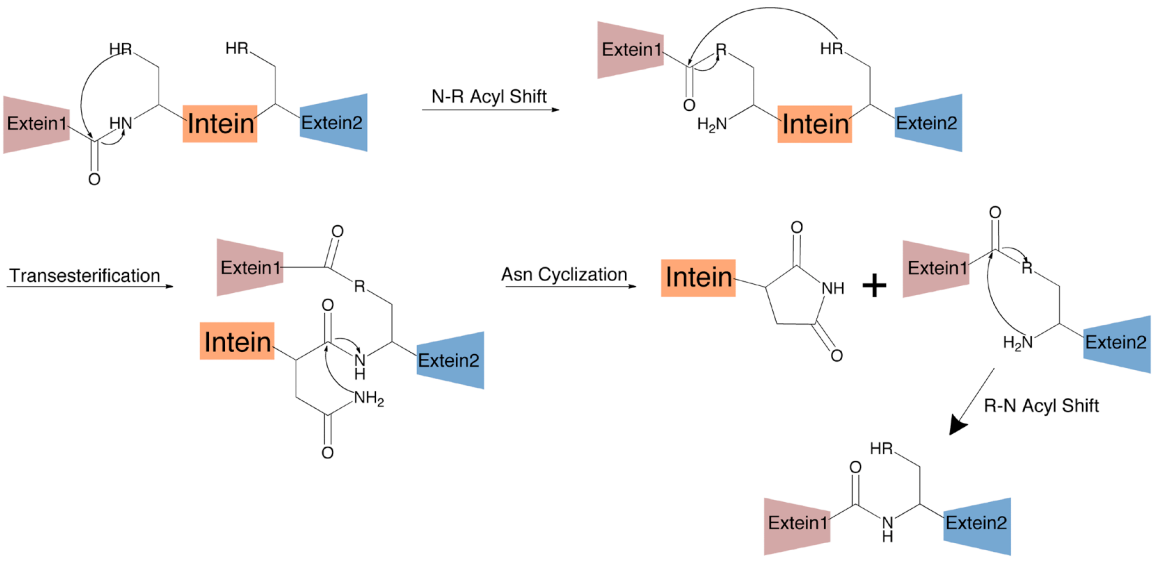


Figure 1.13: Protein splicing using inteins. (R= O, S) Step 1: The Initial N-R acyl rearrangement enables the transesterification in Step 2, splicing the two exteins. Step 3: Intramolecular cyclization of the C-terminal Asn residue of the intein causes its release. The spliced protein then undergoes a spontaneous R-N acyl rearrangement to give the final peptide sequence.

Intein-chemistry was later seen as an easy route to make C-terminal thioesters by using recombinant techniques instead of synthetic techniques.[121] However, isolation of C-terminal thioesters from inteins is difficult since they are quickly captured by flanking exteins. Limiting the *Mxe* intein's ability to capture the C-terminal thioester was performed. By eliminating the intein's C-terminal activity through mutation, the intein was arrested with the thioester in place for capture using an external thiol catalyst *via* transthioesterifcation.[126]

This procedure was quickly marketed by New England Biolabs (NEB) as the Intein Mediated Purification with an Affinity Chitin-binding Tag (IMPACT) system and initially used as a procedure for protein purification.[126] When the protein of interest was fused to the modified intein and a chitin-binding domain (CBD) and passed down a column of chitin resin, the protein-intein-CBD fusion could be isolated. The addition of

an external thiol catalyst then forces transthioesterification, releasing the protein-thioester. Subsequent hydrolysis of the thioester to make the free acid is performed to obtain the purified protein of interest. This new type of affinity chromatography quickly demonstrated superiority for the isolation of protein-thioesters over chemical synthesis while also allowing for longer protein-thioester sequences than those obtained using NCL. Moreover, the technique serves as a method for protein purification. The protein-thioester or the protein-acid could be attained from the same protein expression by altering the external thiol used in the transthioesterification.

Extending the principles of intein splicing past protein purification to perform EPL was the next logical step.[120,124] EPL is a modified version of NCL that uses recombinant expression to obtain the thioester instead of using the synthetic methodologies found in NCL. In EPL, formation of the thioester is achieved by using a modified *Mxe* intein in the presence of an external thiol to facilitate transthioesterification (Figure 1.14). Similar to *in vivo* protein splicing, the process begins with the shift of equilibrium causing an initial N-S acyl rearrangement, making a transient thioester between the fusion protein and the intein (Figure 1.14, Step 1). An external thiol catalyst is added to mediate the transthioesterification with the protein of interest, creating a protein-thioester and leaving the intein bound to the column. However, choice of the correct external thiol catalyst is essential. Thiols such as dithiothreitol (DTT) can cause cyclization and the formation of protein-acids through rapid hydrolysis of the thioester. Attractive thiols for EPL are those that are sufficiently nucleophilic for the initial transthioesterification event, but are also labile enough to be displaced by a second

transthioesterification with a synthetic peptide containing an N-terminal Cys, not by water to form the protein-acid.[121]

Following elution from the column, the protein-thioester is isolated in high yields. When a peptide containing the desired nonproteinogenic amino acid(s) and an N-terminal Cys residue is added to the protein-thioester, another transthioesterification is mediated creating a new transient protein-thioester resulting in the FL protein (Figure 1.14, Step 3). A spontaneous S-N acyl rearrangement produces the FL protein with a native peptide bond (Figure 1.14, Step 4).

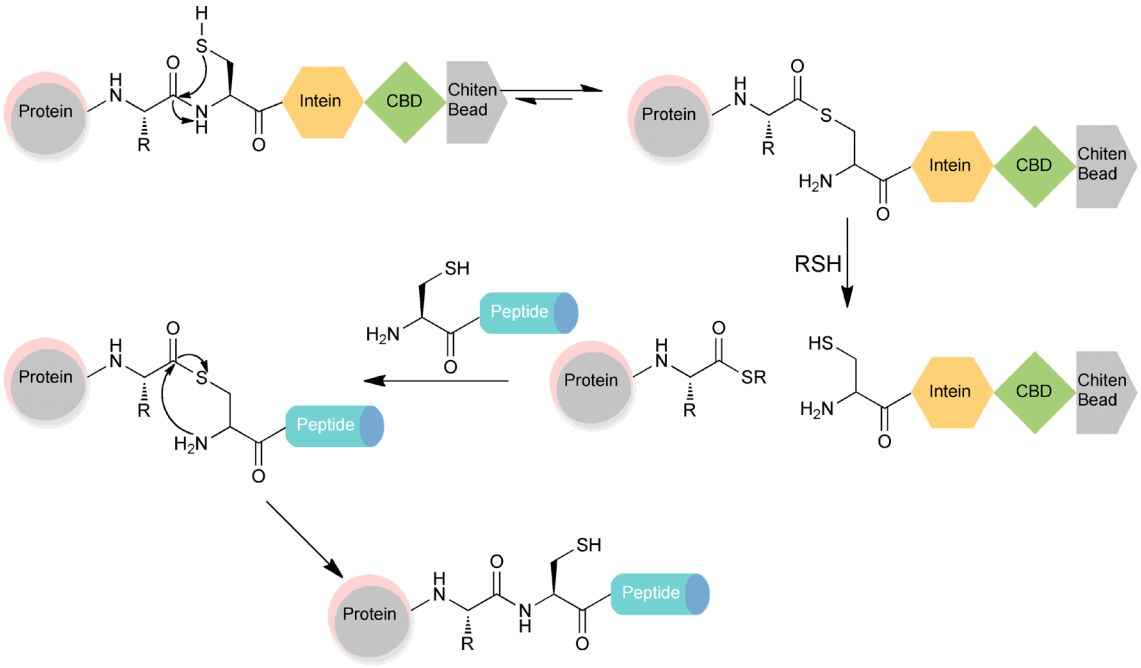


Figure 1.14: General scheme for expressed protein ligation (EPL). Step 1: Equilibrium between the fusion protein Cys and thioester results N-S acyl rearrangement. Step 2: Addition of an external thiol causes transthioesterification between the protein and the thiol giving a protein thioester free from the chitin resin bound remainder of the fusion protein. Step 3: Addition of a synthetic peptide with a N-terminal Cys causes a second transthioesterification followed by a spontaneous S-N acyl rearrangement (Step 4) to give the final FL protein with a peptide bond.

Though EPL can make the isolation of the thioester easier and can greatly increase protein-thioester yields compared to many of the synthetic methodologies currently available, EPL, like NCL, requires the obligate need of a Cys residue at the ligation site. The length of peptide containing the nonproteinogenic amino acid(s) is limited to the lengths available when using SPPS.

1.9 Incorporating Nonproteinogenic Amino Acids Into Azurin and Cu_A Azurin

While the biosynthetic methodologies for incorporation of nonproteinogenic amino acids are appealing, we decided to pursue the incorporation of nonproteinogenic amino acids into the azurin scaffold by using synthetic routes. Azurin and Cu_A azurin (because it is built from the azurin scaffold) are well suited for the use of EPL to incorporate nonproteinogenic amino acids (Figure 1.15). At a length of only 128 residues for azurin (130 residues for Cu_A Az), the protein is small. The convenient accessibility of three of the five copper binding ligands in the C-terminal 17 amino acids of azurin (Cys112, His117, and Met121), and six of the eight copper binding ligands in the C-terminal 19 amino acids in Cu_A azurin (Cys112, Glu114, Cys112, His120, and Met123), provides numerous site-specific locations for incorporating nonproteinogenic amino acids to investigate the specific functionality of the copper ligands. Additionally, the necessary synthesized peptides contain Cys112 at the N-terminus that participates in transthioesterification, affording an effective route to semi-synthesize azurin and Cu_A azurin variants containing nonproteinogenic amino acids by EPL.

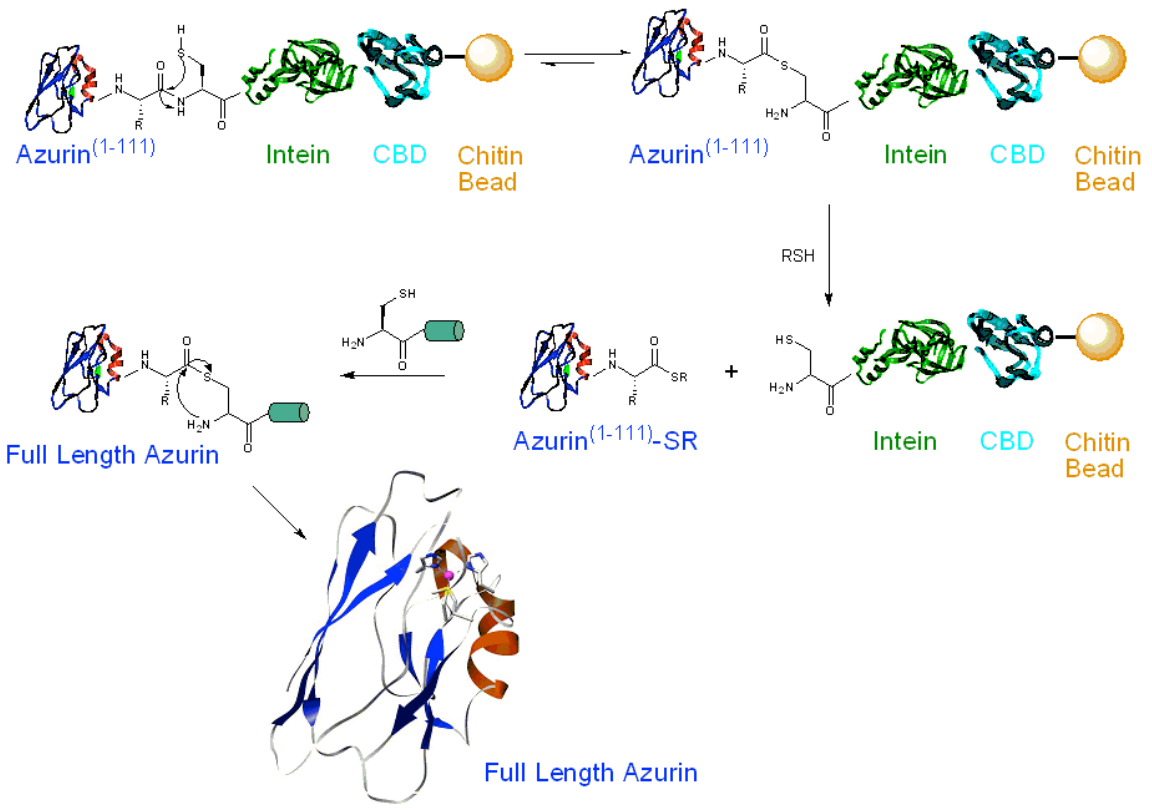


Figure 1.15: Expressed protein ligation of azurin. The first 111 amino acids of azurin are expressed in-frame to the Mxe intein and the chitin-binding domain (CBD) using the pTXB1 vector. After purification by affinity chromatography with chitin beads, the formation of azurin(1-111)-SR is mediated by the intein with the addition of an external thiol. Capture of the azurin-SR is done with the C-terminal 17 amino acid synthetic peptide which after a S-N acyl rearrangement gives FL WT azurin.

To semi-synthesize azurin using EPL, the N-terminal 111 amino acid protein-thioester needed for transthioesterification was made recombinantly using an intein-fusion. Cloning work done by Berry, et al. resulted in an azurin$^{1\text{-}111}$-intein-CBD fusion protein using the NEB pTXB1 vector containing the modified *Mxe* intein.[16] In initial studies, intein-mediated protein cleavage to attain the azurin-SR (protein-thioester) was successful though the azurin-SR readily hydrolyzed in buffer. In an attempt to limit hydrolysis and optimize the azurin-SR stability, ligations at a range of pH were completed but no improvement in yield was observed. Instead of azurin-SR isolation,

attempts to capture the azurin-SR with the synthesized peptide using an on-column cleavage-capture were tried. Though protein yields increased using this methodology, hydrolysis of the thioester remained problematic at pH ranges from 5-8.5. While the yields are far from theoretical yields due to hydrolysis, significant protein yields are obtainable (~5 mg/L growth).

1.10 SUMMARY

While site-directed mutagenesis has remained the primary means for probing ligand functionality in metalloproteins, a new area of investigation has recently began using nonproteinogenic amino acids to more narrowly define ligand function and fine tune metalloproteins. The two metalloproteins that we have chosen to probe using nonproteinogenic amino acid substitution are azurin and Cu_A azurin, two cupredoxins involved in electron transfer. Both proteins exhibit characteristic LMCT bands that produce deep colors (blue in azurin and purple in Cu_A azurin) and distinctive EPR signals. Both proteins are also well studied and changes to either the CT bands or EPR signals reliably show perturbations of electronic and geometric components of the site.

Though many routes for the incorporation of nonproteinogenic amino acids exist, both biosynthetically and synthetically, expressed protein ligation serves as the most straightforward methodology to study both the type 1 copper site in azurin and the Cu_A site in Cu_A azurin by nonproteinogenic amino acid substitution. Since the azurin-thioester protein is readily accessible through recombinant expression and the synthesis of a 17-mer peptide (19-mer peptide for Cu_A Az) containing an N-terminal Cys residue can be achieved, using EPL to incorporate nonproteinogenic amino acids into both scaffolds can

be accomplished. Herein will be described the efforts at probing the type 1 copper site with the nonproteinogenic amino acids selenocysteine (Chapter 2) and homocysteine (Chapters 3 and 4). Furthermore, efforts at understanding the unique bridging thiolates used to bind Cu in Cu_A azurin (Chapter 5) as well as the unique binding properties of the backbone carbonyl from Glu114 in Cu_A azurin (Chapter 6) with nonproteinogenic amino acids will be described.

1.11 REFERENCES

(1) *CRC Practical Handbook of Chemistry and Physics;* 85th ed; CRC Press; **2005**.

(2) Lutgens, F. K.; Tarbuck, E. J. *Essentials of Geology;* 7th ed; Prentice Hall; **2000**.

(3) Audesirk, T.; Audesirk, G. *Biology, Life on Earth;* 5th ed; Prentice Hall; **1999**.

(4) Enger, E. D.; Ross, F. C. *Concepts in Biology;* 10th ed; McGraw-Hill; **2003**.

(5) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* "Structural and Functional Aspects of Metal Sites in Biology." **1996**, *96*, 2239-2314.

(6) Uauy, R.; Olivares, M.; Gonzalez, M. *Am J Clin Nutr.* "Essentiality of copper in humans." **1998**, *67*, 952S-959.

(7) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. *Chem. Rev.* "Electronic Structures of Active Sites in Copper Proteins: Contributions to Reactivity." **1992**, *92*, 521-542.

(8) Bento, I.; Carrondo, M.; Lindley, P. *J. Biol. Inorg. Chem.* "Reduction of dioxygen by enzymes containing copper." **2006**, *11*, 539-547.

(9) Claus, H.; Decker, H. *Syst. Appl. Microbiol.* "Bacterial tyrosinases." **2006**, *29*, 3-14.

(10) Halaouli, S.; Asther, M.; Sigoillot, J. C.; Hamdi, M.; Lomascolo, A. *J. Appl. Microbiol.* "Fungal tyrosinases: new prospects in molecular characteristics, bioengineering and biotechnological applications." **2006**, *100*, 219-232.

(11) Harrison, M. D.; Jones, C. E.; Dameron, C. T. *J. Biol. Inorg. Chem.* "Copper chaperones: function, structure and copper-binding properties." **1999**, *4*, 145-153.

(12) MacPherson, I.; Murphy, M. *Cell. Mol. Life Sci.* "Type-2 copper-containing enzymes." **2007**, *64*, 2887-2899.

(13) McGuirl, M. A.; Dooley, D. M. *Curr. Opin. Chem. Biol.* "Copper-containing oxidases." **1999**, *3*, 138-144.

(14) Torres, J.; Wilson, M. T. *Biochim. Biophys. Acta, Bioenerg.* "The reactions of copper proteins with nitric oxide." **1999**, *1411*, 310-322.

(15) Messerschmidt, A.; Rossi, A.; Ladenstein, R.; Huber, R.; Bolognesi, M.; Gatti, G.; Marchesini, A.; Petruzzelli, R.; Finazzi-Agru, A. *J. Mol. Biol.* "X-ray crystal structure of the blue oxidase ascorbate oxidase from Zucchini : Analysis of the polypeptide fold and a model of the copper sites and ligands." **1989**, *206*, 513-529.

(16) Berry, S. M.; Ralle, M.; Low, D. W.; Blackburn, N. J.; Lu, Y. *J. Am. Chem. Soc.* "Probing the Role of Axial Methionine in the Blue Copper Center of Azurin with Unnatural Amino Acids." **2003**, *125*, 8760-8768.

(17) Garner, D. K.; Vaughan, M. D.; Hwang, H. J.; Savelieff, M. G.; Berry, S. M.; Honek, J. F.; Lu, Y. *J. Am. Chem. Soc.* "Reduction Potential Tuning of the Blue Copper Center in *Pseudomonas aeruginosa* Azurin by the Axial Methionine as Probed by Unnatural Amino Acids." **2006**, *128*, 15608-15617.

(18) Malmström, B. G. *Eur. J. Biochem.* "Rack-Induced Bonding In Blue-Copper Proteins." **1994**, *223*, 711-718.

(19) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Roos, B. O. *J. Am. Chem. Soc.* "Theoretical Study of the Electronic Spectrum of Plastocyanin." **1997**, *119*, 218-226.

(20) Andrew, C. R.; Yeom, H.; Valentine, J. S.; Karlsson, B. G.; van Pouderoyen, G.; Canters, G. W.; Loehr, T. M.; Sanders-Loehr, J.; Bonander, N. *J. Am. Chem. Soc.* "Raman Spectroscopy as an Indicator of Cu-S Bond Length in Type 1 and Type 2 Copper Cysteinate Proteins." **1994**, *116*, 11489-98.

(21) DeBeer George, S.; Basumallick, L.; Szilagyi, R. K.; Randall, D. W.; Hill, M. G.; Nersissian, A. M.; Valentine, J. S.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* "Spectroscopic Investigation of Stellacyanin Mutants: Axial Ligand Interactions at the Blue Copper Site." **2003**, *125*, 11314-11328.

(22) den Blaauwen, T.; Canters, G. W. *J. Am. Chem. Soc.* "Creation of type-1 and type-2 copper sites by addition of exogenous ligands to the *Pseudomonas aeruginosa* azurin His117Gly mutant." **1993**, *115*, 1121-1129.

(23) Gray, H. B.; Malmström, B. G.; Williams, R. J. P. *J. Biol. Inorg. Chem.* "Copper Coordination in Blue Proteins." **2000**, *5*, 551-559.

(24) Guckert, J. A.; Lowery, M. D.; Solomon , E. I. *J. Am. Chem. Soc.* "Electronic Structure of the Reduced Blue Copper Active Site: Contributions to Reduction Potentials and Geometry." **1995**, *117*, 2817-2844.

(25) Kofman, V.; Farver, O.; Pecht, I.; Goldfarb, D. *J. Am. Chem. Soc.* "Two-Dimensional Pulsed EPR Spectroscopy of the Copper Protein Azurin." **1996**, *118*, 1201-1206.

(26) Murphy, L. M.; Strange, R. W.; Karlsson, B. G.; Lundberg, L. G.; Pascher, T.; Reinhammar, B.; Hasnain, S. S. *Biochemistry.* "Structural characterization of azurin from *Pseudomonas aeruginosa* and some of its methionine-121 mutants." **1993**, *32*, 1965-1975.

(27) Olsson, M. H. M.; Ryde, U. *J. Biol. Inorg. Chem.* "The influence of axial ligands on the reduction potential of blue copper proteins." **1999**, *4*, 654-663.

(28) Solomon, E. I.; Hedman, B.; Hodgson, K. O.; Dey, A.; Szilagyi, R. K. *Coord. Chem. Rev.* "Ligand K-edge X-ray absorption spectroscopy: covalency of ligand-metal bonds." **2005**, *249*, 97-129.

(29) Solomon, E. I.; Penfield, K. W.; Gewirth, A. A.; Lowery, M. D.; Shadle, S. E.; Guckert, J. A.; LaCroix, L. B. *Inorg. Chim. Acta.* "Electronic structure of the oxidized and reduced blue copper sites: contributions to the electron transfer pathway, reduction potential, and geometry." **1996**, *243*, 67-78.

(30) van Pouderoyen, G.; Andrew, C. R.; Loehr, T. M.; Sanders-Loehr, J.; Mazumdar, S.; Hill, H. A. O.; Canters, G. W. *Biochemistry.* "Spectroscopic and Mechanistic Studies of Type-1 and Type-2 Copper Sites in *Pseudomonas aeruginosa*

Azurin As Obtained by Addition of External Ligands to Mutant His46Gly." **1996**, *35*, 1397-1407.

(31) Hough, M. A.; Strange, R. W.; Hasnain, S. S. *J. Mol. Biol.* "Conformational variability of the Cu site in one subunit of bovine CuZn superoxide dismutase: the importance of mobility in the Glu119-Leu142 loop region for catalytic function." **2000**, *304*, 231-241.

(32) Karlsson, B. G.; Tsai, L.-C.; Nar, H.; Sanders-Loehr, J.; Bonander, N.; Langer, V.; Sjolin, L. *Biochemistry.* "X-ray Structure Determination and Characterization of the *Pseudomonas aeruginosa* Azurin Mutant Met121Glu." **1997**, *36*, 4089-4095.

(33) Nar, H.; Messerschmidt, A.; Huber, R.; van de Kamp, M.; Canters, G. W. *J. Mol. Biol.* "Crystal structure analysis of oxidized *Pseudomonas aeruginosa* azurin at pH 5.5 and pH 9.0 : A pH-induced conformational transition involves a peptide bond flip." **1991**, *221*, 765-772.

(34) Romero, A.; Hoitink, C. W. G.; Nar, H.; Huber, R.; Messerschmidt, A.; Canters, G. W. *J. Mol. Biol.* "X-ray Analysis and Spectroscopic Characterization of M121Q Azurin : A Copper Site Model for Stellacyanin." **1993**, *229*, 1007-1021.

(35) Yamakura, F.; Sugio, S.; Hiraoka, B. Y.; Ohmori, D.; Yokota, T. *Biochemistry.* "Pronounced Conversion of the Metal-Specific Activity of Superoxide Dismutase from *Porphyromonas gingivalis* by the Mutation of a Single Amino Acid (Gly155Thr) Located Apart from the Active Site." **2003**, *42*, 10790-10799.

(36) Chang, T. K.; Iverson, S. A.; Rodrigues, C. G.; Kiser, C. N.; Lew, A. Y. C.; Germanas, J. P.; Richards, J. H. *Proc. Natl. Acad. Sci. USA.* "Gene Synthesis, Expression, and Mutagenesis Of the Blue Copper Proteins Azurin and Plastocyanin." **1991**, *88*, 1325-1329.

(37) den Blaauwen, T.; van de Kamp, M.; Canters, G. W. *J. Am. Chem. Soc.* "Type I and II copper sites obtained by external addition of ligands to a His117Gly azurin mutant." **1991**, *113*, 5050-5052.

(38) Germanas, J. P.; Di Bilio, A. J.; Gray, H. B.; Richards, J. H. *Biochemistry.* "Site saturation of the histidine-46 position in *Pseudomonas aeruginosa* azurin: Characterization of the His46Asp copper and cobalt proteins." **1993**, *32*, 7698-7702.

(39) Karlsson, B. G.; Aasa, R.; Malmstrom, B. G.; Lundberg, L. G. *FEBS Lett.* "Rack-induced bonding in blue copper proteins: Spectroscopic properties and reduction potential of the azurin mutant Met-121 --> Leu." **1989**, *253*, 99-102.

(40) Mizoguchi, T. J.; Di Bilio, A. J.; Gray, H. B.; Richards, J. H. *J. Am. Chem. Soc.* "Blue to type 2 binding. Copper(II) and cobalt(II) derivatives of a Cys112Asp mutant of *Pseudomonas aeruginosa* azurin." **1992**, *114*, 10076-10078.

(41) Faham, S.; Mizoguchi, T. J.; Adman, E. T.; Gray, H. B.; Richards, J. H.; Rees, D. C. *J. Biol. Inorg. Chem.* "Role of the active-site cysteine of *Pseudomonas aeruginosa* azurin. Crystal structure analysis of the CuII(Cys112Asp) protein." **1997**, *2*, 464-469.

(42) DeBeer, S.; Kiser, C. N.; Mines, G. A.; Richards, J. H.; Gray, H. B.; Solomon, E. I.; Hedman, B.; Hodgson, K. O. *Inorg. Chem.* "X-ray Absorption Spectra of the Oxidized and Reduced Forms of C112D Azurin from *Pseudomonas aeruginosa*." **1999**, *38*, 433-438.

(43) Czernuszewicz, R. S.; Fraczkiewicz, G.; Zareba, A. A. *Inorg. Chem.* "A Detailed Resonance Raman Spectrum of Nickel(II)-Substituted *Pseudomonas aeruginosa* Azurin." **2005**, *44*, 5745-5752.

(44) den Blaauwen, T.; Hoitink, C. W. G.; Canters, G. W.; Han, J.; Loehr, T. M.; Sanders-Loehr, J. *Biochemistry.* "Resonance Raman Spectroscopy of the Azurin His117Gly Mutant. Interconversion of Type 1 and Type 2 Copper Sites Through Exogenous Ligands." **1993**, *32*, 12455-12464.

(45) Funk, T.; Kennepohl, P.; Di Bilio, A. J.; Wehbi, W. A.; Young, A. T.; Friedrich, S.; Arenholz, E.; Gray, H. B.; Cramer, S. P. *J. Am. Chem. Soc.* "X-ray Magnetic Circular Dichroism of *Pseudomonas aeruginosa* Nickel(II) Azurin." **2004**, *126*, 5859-5866.

(46) Shafaat, H. S.; Leigh, B. S.; Tauber, M. J.; Kim, J. E. *J. Phys. Chem. B.* "Resonance Raman Characterization of a Stable Tryptophan Radical in an Azurin Mutant." **2009**, *113*, 382-388.

(47) Woodruff, W. H.; Norton, K. A.; Swanson, B. I.; Fry, H. A. *Proc. Natl. Acad. Sci. USA.* "Temperature Dependence of the Resonance Raman Spectra of

Plastocyanin and Azurin Between Cryogenic and Ambient Conditions." **1984**, *81*, 1263-1267.

(48) Fittipaldi, M.; Warmerdam, G. C. M.; de Waal, E. C.; Canters, G. W.; Cavazzini, D.; Rossi, G. L.; Huber, M.; Groenen, E. J. J. *ChemPhysChem.* "Spin-Density Distribution in the Copper Site of Azurin." **2006**, *7*, 1286-1293.

(49) Sarangi, R.; Gorelsky, S. I.; Basumallick, L.; Hwang, H. J.; Pratt, R. C.; Stack, T. D. P.; Lu, Y.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *J. Am. Chem. Soc.* "Spectroscopic and Density Functional Theory Studies of the Blue-Copper Site in M121SeM and C112SeC Azurin: Cu-Se Versus Cu-S Bonding." **2008**, *130*, 3866-3877.

(50) Nasim, M. T.; Jaenecke, S.; Belduz, A.; Kollmus, H.; Flohe, L.; McCarthy, J. E. G. *J. Biol. Chem.* "Eukaryotic Selenocysteine Incorporation Follows a Nonprocessive Mechanism That Competes with Translational Termination." **2000**, *275*, 14846-14852.

(51) Frank, P.; Licht, A.; Tullius, T. D.; Hodgson, K. O.; Pecht, I. *J. Biol. Chem.* "A selenomethionine-containing azurin from an auxotroph of *Pseudomonas aeruginosa*." **1985**, *260*, 5518-5525.

(52) Karlsson, B. G.; Nordling, M.; Pascher, T.; Tsai, L.-C.; Sjolin, L.; Lundberg, L. G. *Protein Eng.* "Cassette mutagenesis of Met121 in azurin from *Pseudomonas aeruginosa*." **1991**, *4*, 343-349.

(53) Lu, Y. *Cupredoxins in Comprehensive Coordination Chemistry II: From Biology to Nanotechnology;* ed; Elsevier; **2003**.

(54) Solomon, E. I.; Randall, D. W.; Glaser, T. *Coord. Chem. Rev.* "Electronic structures of active sites in electron transfer metalloproteins: contributions to reactivity." **2000**, *200-202*, 595-632.

(55) Crane, B. R.; Di Bilio, A. J.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* "Electron Tunneling in Single Crystals of *Pseudomonas aeruginosa* Azurins." **2001**, *123*, 11623-11631.

(56) Li, H.; Webb, S. P.; Ivanic, J.; Jensen, J. H. *J. Am. Chem. Soc.* "Determinants of the Relative Reduction Potentials of Type-1 Copper Sites in Proteins." **2004**, *126*, 8010-8019.

(57) Di Bilio, A. J.; Chang, T. K.; Malmström, B. G.; Gray, H. B.; Karlsson, B. G.; Nordling, M.; Pascher, T.; Lundberg, L. G. *Inorg. Chim. Acta.* "Electronic absorption spectra of M(II)(Met121X) azurins (M=Co, Ni, Cu; X=Leu, Gly, Asp, Glu): charge-transfer energies and reduction potentials." **1992**, *198-200*, 145-148.

(58) Kroes, S. J.; Hoitink, C. W. G.; Andrew, C. R.; Ai, J.; Sanders-Loehr, J.; Messerschmidt, A.; Hagen, W. R.; Canters, G. W. *Eur. J. Biochem.* "The Mutation Met121His Creates a Type-1.5 Copper Site in *Alcaligenes denitrificans* Azurin." **1996**, *240*, 342-351.

(59) Pascher, T.; Karlsson, B. G.; Nordling, M.; Malmström, B. G.; Vaenngaard, T. *Eur. J. Biochem.* "Reduction potentials and their pH dependence in site-directed-mutant forms of azurin from *Pseudomonas aeruginosa*." **1993**, *212*, 289-296.

(60) Diederix, R. E. M.; Canters, G. W.; Dennison, C. *Biochemistry.* "The Met99Gln Mutant of Amicyanin from *Paracoccus versutus*." **2000**, *39*, 9551-9560.

(61) Harrison, M. D.; Dennison, C. *ChemBioChem.* "An Axial Met Ligand at a Type 1 Copper Site is Preferable for Fast Electron Transfer." **2004**, *5*, 1579-1581.

(62) Hough, M. A.; Ellis, M. J.; Antonyuk, S.; Strange, R. W.; Sawers, G.; Eady, R. R.; Samar Hasnain, S. *J. Mol. Biol.* "High Resolution Structural Studies of Mutants Provide Insights into Catalysis and Electron Transfer Processes in Copper Nitrite Reductase." **2005**, *350*, 300-309.

(63) Olesen, K.; Veselov, A.; Zhao, Y.; Wang, Y.; Danner, B.; Scholes, C. P.; Shapleigh, J. P. *Biochemistry.* "Spectroscopic, Kinetic, and Electrochemical Characterization of Heterologously Expressed Wild-Type and Mutant Forms of Copper-Containing Nitrite Reductase from *Rhodobacter sphaeroides*." **1998**, *37*, 6086-6094.

(64) Xu, F.; Palmer, A. E.; Yaver, D. S.; Berka, R. M.; Gambetta, G. A.; Brown, S. H.; Solomon, E. I. *J. Biol. Chem.* "Targeted Mutations in a Trametes villosa Laccase. Axial Perturbations of the T1 Copper." **1999**, *274*, 12372-12375.

(65) Bauer, R.; Danielsen, E.; Hemmingsen, L.; Bjerrum, M. J.; Hansson, O.; Singh, K. *J. Am. Chem. Soc.* "Interplay between Oxidation State and Coordination Geometry of Metal Ions in Azurin." **1997**, *119*, 157-162.

(66) Farver, O.; Skov, L. K.; Pascher, T.; Karlsson, B. G.; Nordling, M.; Lundberg, L. G.; Vaenngaard, T.; Pecht, I. *Biochemistry.* "Intramolecular electron transfer in single-site-mutated azurins." **1993**, *32*, 7317-7322.

(67) Salgado, J.; Kroes, S. J.; Berg, A.; Moratal, J. M.; Canters, G. W. *J. Biol. Chem.* "The Dynamic Properties of the M121H Azurin Metal Site as Studied by NMR of the Paramagnetic Cu(II) and Co(II) Metalloderivatives." **1998**, *273*, 177-185.

(68) Han, J.; Loehr, T. M.; Lu, Y.; Valentine, J. S.; Averill, B. A.; Sanders-Loehr, J. *J. Am. Chem. Soc.* "Resonance Raman excitation profiles indicate multiple Cys-Cu charge transfer transitions in type 1 copper proteins." **1993**, *115*, 4256-4263.

(69) LaCroix, L. B.; Randall, D. W.; Nersissian, A. M.; Hoitink, C. W. G.; Canters, G. W.; Valentine, J. S.; Solomon, E. I. *J. Am. Chem. Soc.* "Spectroscopic and Geometric Variations in Perturbed Blue Copper Centers: Electronic Structures of Stellacyanin and Cucumber Basic Protein." **1998**, *120*, 9621-9631.

(70) LaCroix, L. B.; Shadle, S. E.; Wang, Y.; Averill, B. A.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* "Electronic Structure of the Perturbed Blue Copper Site in Nitrite Reductase: Spectroscopic Properties, Bonding, and Implications for the Entatic/Rack State." **1996**, *118*, 7755-7768.

(71) Messerschmidt, A.; Prade, L.; Kroes, S. J.; Sanders-Loehr, J.; Huber, R.; Canters, G. W. *Proc. Natl. Acad. Sci. USA.* "Rack-induced metal binding vs. flexibility: Met121His azurin crystal structures at different pH." **1998**, *95*, 3443-3448.

(72) Olsson, M. H. M.; Ryde, U. *J. Am. Chem. Soc.* "Geometry, Reduction Potential, and Reorganization Energy of the Binuclear CuA Site, Studied by Density Functional Theory." **2001**, *123*, 7866-7876.

(73) DeBeer George, S.; Metz, M.; Szilagyi, R. K.; Wang, H.; Cramer, S. P.; Lu, Y.; Tolman, W. B.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* "A Quantitative Description of the Ground-State Wave Function of CuA by X-ray Absorption Spectroscopy; Comparison to Plastocyanin and Relevance to Electron Transfer." **2001**, *123*, 5757-5767.

(74) Farver, O.; Hwang, H. J.; Lu, Y.; Pecht, I. *J. Phys. Chem. B.* "Reorganization Energy of the CuA Center in Purple Azurin: Impact of the Mixed Valence-to-Trapped Valence State Transition." **2007**, *111*, 6690-6694.

(75) Farver, O.; Lu, Y.; Ang, M. C.; Pecht, I. *Proc. Natl. Acad. Sci. USA.* "Enhanced rate of intramolecular electron transfer in an engineered purple CuA azurin." **1999**, *96*, 899-902.

(76) Robinson, H.; Ang, M. C.; Gao, Y.-G.; Hay, M. T.; Lu, Y.; Wang, A. H. J. *Biochemistry.* "Structural Basis of Electron Transfer Modulation in the Purple Cu_A Center." **1999**, *38*, 5677-5683.

(77) Sun, D.; Wang, X.; Davidson, V. L. *Arch. Biochem. Biophys.* "Redox properties of an engineered purple CuA azurin." **2002**, *404*, 158-162.

(78) Hwang, H. J.; Berry, S. M.; Nilges, M. J.; Lu, Y. *J. Am. Chem. Soc.* "Axial Methionine Has Much Less Influence on Reduction Potentials in a CuA Center than in a Blue Copper Center." **2005**, *127*, 7274-7275.

(79) Houser, R. P.; Halfen, J. A.; Young, V. G.; Blackburn, N. J.; Tolman, W. B. *J. Am. Chem. Soc.* "Structural Characterization of the First Example of a Bis(mu-thiolato)dicopper(II) Complex. Relevance to Proposals for the Electron Transfer Sites in Cytochrome c Oxidase and Nitrous Oxide Reductase." **1995**, *117*, 10745-10746.

(80) So, H. *Bull. Korean Chem. Soc.* "Anaylsis of Intramolecular Electron Transfer in a Mixed-Valence Cu(I)-Cu(II) Complex Using the PKS Model." **1992**, *13*, 385-387.

(81) Shinzawa-Itoh, K.; Aoyama, H.; Muramoto, K.; Terada, H.; Kurauchi, T.; Tadehara, Y.; Yamasaki, A.; Sugimura, T.; Kurono, S.; Tsujimoto, K.; Mizushima, T.; Yamashita, E.; Tsukihara, T.; Yoshikawa, S. *EMBO J.* "Structures and physiological roles of 13 integral lipids of bovine heart cytochrome c oxidase." **2007**, *26*, 1713-1725.

(82) Brown, K.; Djinovic-Carugo, K.; Haltia, T.; Cabrito, I.; Saraste, M.; Moura, J. J. G.; Moura, I.; Tegoni, M.; Cambillau, C. *J. Biol. Chem.* "Revisiting the Catalytic CuZ Cluster of Nitrous Oxide (N2O) Reductase." **2000**, *275*, 41133-41136.

(83) Hay, M.; Richards, J. H.; Lu, Y. *Proc. Natl. Acad. Sci. USA.* "Construction and characterization of an azurin analog for the purple copper site in cytochrome *c* oxidase." **1996**, *93*, 461-464.

(84) Hay, M. T. Ph.D., University of Illinois at Urbana-Champaign, 1998.

(85) Hay, M. T.; Ang, M. C.; Gamelin, D. R.; Solomon, E. I.; Antholine, W. E.; Ralle, M.; Blackburn, N. J.; Massey, P. D.; Wang, X.; Kwon, A. H.; Lu, Y. *Inorg.*

Chem. "Spectroscopic characterization of an engineered purple Cu_A center in azurin." **1998**, *37*, 191-198.

(86) Hwang, H. J.; Nagraj, N.; Lu, Y. *Inorg. Chem.* "Spectroscopic Characterizations of Bridging Cysteine Ligand Variants of an Engineered Cu2(SCys)2 CuA Azurin." **2006**, *45*, 102-107.

(87) Fersht, A. R. *Biochemistry.* "Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase." **1977**, *16*, 1025-1030.

(88) Fersht, A. R.; Dingwall, C. *Biochemistry.* "Evidence for the double-sieve editing mechanism in protein synthesis. Steric exclusion of isoleucine by valyl-tRNA synthetases." **1979**, *18*, 2627-2631.

(89) Rotruck, J. T.; Pope, A. L.; Ganther, H. E.; Swanson, A. B.; Hafeman, D. G.; Hoekstra, W. G. *Science.* "Selenium: Biochemical Role as a Component of Glutathione Peroxidase." **1973**, *179*, 588-590.

(90) Turner, D. C.; Stadtman, T. C. *Arch. Biochem. Biophys.* "Purification of protein components of the clostridial glycine reductase system and characterization of protein A as a selenoprotein." **1973**, *154*, 366-381.

(91) Cone, J. E.; Del Rio, R. M.; Davis, J. N.; Stadtman, T. C. *Proc. Natl. Acad. Sci. USA.* "Chemical characterization of the selenoprotein component of clostridial glycine reductase: identification of selenocysteine as the organoselenium moiety." **1976**, *73*, 2659-2663.

(92) Atkins, J. F.; Gesteland, R. *Science.* "Biochemistry: The 22nd Amino Acid." **2002**, *296*, 1409-1410.

(93) Hao, B.; Gong, W.; Ferguson, T. K.; James, C. M.; Krzycki, J. A.; Chan, M. K. *Science.* "A New UAG-Encoded Residue in the Structure of a Methanogen Methyltransferase." **2002**, *296*, 1462-1466.

(94) Srinivasan, G.; James, C. M.; Krzycki, J. A. *Science.* "Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA." **2002**, *296*, 1459-1462.

(95) Arner, E. S. J.; Sarioglu, H.; Lottspeich, F.; Holmgren, A.; Bock, A. *J. Mol. Biol.* "High-level expression in Escherichia coli of selenocysteine-containing rat

thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the selA, selB and selC genes." **1999**, *292*, 1003-1016.

(96) Leinfelder, W.; Forchhammer, K.; Zinoni, F.; Sawers, G.; Mandrand-Berthelot, M. A.; Bock, A. *J. Bacteriol.* "*Escherichia coli* genes whose products are involved in selenium metabolism." **1988**, *170*, 540-546.

(97) Stadtman, T. C. *Annu. Rev. Biochem.* "Selenocysteine." **1996**, *65*, 83-100.

(98) Sunde, R.; Evenson, J. *J. Biol. Chem.* "Serine incorporation into the selenocysteine moiety of glutathione peroxidase." **1987**, *262*, 933-937.

(99) Feng, L.; Tumbula-Hansen, D.; Min, B.; Namgoong, S.; Salazar, J.; Orellana, O.; Soll, D. In *The Aminoacyl-tRNA Synthetases*; Michael Ibba, C. F., Stephen Cusack, Ed.; Landes Biosciences: Georgetown, TX, 2005, p 314-319.

(100) Hoffmann, P. R.; Berry, M. J. *Thyroid.* "Selenoprotein Synthesis: A Unique Translational Mechanism Used by a Diverse Family of Proteins." **2005**, *15*, 769-775.

(101) Fischer, N.; Paleskava, A.; Gromadski, K. B.; Konevega, A. L.; Wahl, M. C.; Stark, H.; Rodnina, M. V. *Biol. Chem.* "Towards Understanding Selenocysteine Incorporation into Bacterial Proteins." **2007**, *388*, 1061-1067.

(102) Su, D.; Li, Y.; Gladyshev, V. N. *Nucleic Acids Res.* "Selenocysteine insertion directed by the 3'-UTR SECIS element in *Escherichia coli*." **2005**, *33*, 2486-2492.

(103) Kryukov, G. V.; Castellano, S.; Novoselov, S. V.; Lobanov, A. V.; Zehtab, O.; Guigo, R.; Gladyshev, V. N. *Science.* "Characterization of Mammalian Selenoproteomes." **2003**, *300*, 1439-1443.

(104) Aldag, C.; Gromov, I. A.; Garcia-Rubio, I.; von Koenig, K.; Schlichting, I.; Jaun, B.; Hilvert, D. *Proc. Natl. Acad. Sci. USA.* "Probing the role of the proximal heme ligand in cytochrome P450cam by recombinant incorporation of selenocysteine." **2009**, *106*, 5481-5486.

(105) Benzer, S.; Champe, S. P. *Proc. Natl. Acad. Sci. USA.* "A Change from Nonsense to sense in the genetic code." **1962**, *48*, 1114-1121.

(106) Garen, A.; Siddiqui, O. *Proc. Natl. Acad. Sci. USA.* "Supression of Mutations in the Alkaline Phosphatase Structural Cistron of *E. coli*." **1962**, *48*, 1121-1127.

(107) Xie, J.; Schultz, P. G. *Nat. Rev. Mol. Cell Biol.* "A chemical toolkit for proteins - an expanded genetic code." **2006**, *7*, 775-782.

(108) Brustad, E.; Bushey, M. L.; Brock, A.; Chittuluru, J.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* "A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins." **2008**, *18*, 6004-6006.

(109) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chem. Int. Ed.* "A Facile System for Encoding Unnatural Amino Acids in Mammalian Cells." **2009**, *48*, 4052-4055.

(110) Lee, H. S.; Spraggon, G.; Schultz, P. G.; Wang, F. *J. Am. Chem. Soc.* "Genetic Incorporation of a Metal-Ion Chelating Amino Acid into Proteins as a Biophysical Probe." **2009**, *131*, 2481-2483.

(111) Liu, C. C.; Mack, A. V.; Tsao, M.-L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V. *Proc. Natl. Acad. Sci. USA.* "Protein evolution with an expanded genetic code." **2008**, *105*, 17688-17693.

(112) Cheong, J.-J.; Hwang, I.; Rhee, S.; Moon, T.; Choi, Y.; Kwon, H.-B. *Biotechnol. Lett.* "Complementation of an *E. coli* cysteine auxotrophic mutant for the structural modification study of 3′(2′),5′-bisphosphate nucleotidase." **2007**, *29*, 913-918.

(113) Neuwald, A. F.; Krishnan, B. R.; Brikun, I.; Kulakauskas, S.; Suziedelis, K.; Tomcsanyi, T.; Leyh, T. S.; Berg, D. E. *J. Bacteriol.* "cysQ, a gene needed for cysteine synthesis in *Escherichia coli* K-12 only during aerobic growth." **1992**, *174*, 415-425.

(114) Boschi-Muller, S.; Muller, S.; Van Dorsselaer, A.; Bock, A.; Branlant, G. *FEBS Lett.* "Substituting selenocysteine for active site cysteine 149 of phosphorylating glyceraldehyde 3-phosphate dehydrogenase reveals a peroxidase activity." **1998**, *439*, 241-245.

(115) Sabine, M.; Hans, S.; Bernard, G.; Walter, V.; Christian, B.; August, B. *Biochemistry.* "The Formation of Diselenide Bridges in Proteins by Incorporation of

Selenocysteine Residues: Biosynthesis and Characterization of (Se)2-Thioredoxin." **1994**, *33*, 3404-3412.

(116) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. *Science.* "Synthesis of proteins by native chemical ligation." **1994**, *266*, 776-779.

(117) Aimoto, S. *Pept. Sci.* "Polypeptide synthesis by the thioester method." **1999**, *51*, 247-265.

(118) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. USA.* "Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology." **1999**, *96*, 10068-10073.

(119) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* "Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives." **1997**, *119*, 4325-4329.

(120) Muir, T. W. *Annu. Rev. Biochem.* "Semisynthesis of proteins by Expressed Protein Ligation." **2003**, *72*, 249-289.

(121) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* "Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction." **2006**, *128*, 6640-6646.

(122) Pentelute, B. L.; Kent, S. B. H. *Org. Lett.* "Selective Desulfurization of Cysteine in the Presence of Cys(Acm) in Polypeptides Obtained by Native Chemical Ligation." **2007**, *9*, 687-690.

(123) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* "Staudinger Ligation: A Peptide from a Thioester and Azide." **2000**, *2*, 1939-1941.

(124) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. USA.* "Expressed protein ligation: A general method for protein engineering." **1998**, *95*, 6705-6710.

(125) Muralidharan, V.; Muir, T. W. *Nat. Meth.* "Protein ligation: an enabling technology for the biophysical analysis of proteins." **2006**, *3*, 429-438.

(126) Chong, S.; Mersha, F. B.; Comb, D. G.; Scott, M. E.; Landry, D.; Vence, L. M.; Perler, F. B.; Benner, J.; Kucera, R. B.; Hirvonen, C. A. *Gene.* "Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element." **1997**, *192*, 271-281.

(127) David, R.; Richter, M. P. O.; Beck-Sickinger, A. G. *Eur. J. Biochem.* "Expressed protein ligation Method and applications." **2004**, *271*, 663-677.

(128) Perler, F. B.; Adam, E. *Curr. Opin. Biotechnol.* "Protein splicing and its applications." **2000**, *11*, 377-383.

(129) Shao, Y.; Xu, M.-Q.; Paulus, H. *Biochemistry.* "Protein Splicing: Evidence for an N-O Acyl Rearrangement as the Initial Step in the Splicing Process." **1996**, *35*, 3810-3815.

CHAPTER 2

INVESTIGATING THE RACK HYPOTHESIS: CYS112HCY AZURIN

2.1 INTRODUCTION

Introducing functional groups that are unavailable in the 20 proteinogenic amino acids by incorporating nonproteinogenic amino acids can fine tune function and define metal ligand functionality.[1-5] Moreover, unique functional groups can act as an initial step in metalloprotein engineering.[5-7] *Pseudomonas aeurginosa* blue copper azurin has been found to be a very good metalloprotein model system for the incorporation of nonproteinogenic amino acids.[1-5] The blue copper or type 1 copper site in azurin has trigonal-bipyramidal geometry with the His_2Cys equatorial plane consisting of His46, His117, and Cys112. The unique bipyramidal geometry is created in azurin by the presence of two axial ligands interactions. Typical of many type 1 copper sites, a Met residue (Met121) acts as an axial ligand. In addition, the azurin type 1 copper site has an anionic interaction with the backbone carbonyl oxygen from Gly45 (See Figure 1.4).

The type 1 copper proteins display unique spectroscopic properties (strong ligand-to-copper charge transfer bands and small parallel hyperfine splitting in EPR) and increased redox potentials (300-600 mV) that are not found in most copper complexes outside the blue copper scaffold (See Figure 1.5). These unique characteristics have been attributed to the so-called "rack hypothesis" which states that the unique properties inferred to the blue copper proteins (including azurin) are a direct function of the entatic geometry enforced by the azurin protein scaffold.[8-12] The constrained geometry reduces the Jahn-Teller effect normally seen in copper complexes and thus reduces the

reorganization energy required during electron transfer, making blue copper proteins efficient electron transfer agents.[13-15] The rack hypothesis has been supported by extensive investigation by spectroscopic and crystallographic methods.[8-12,16,17] These studies also defined the roles of the metal ligands in azurin. Consequently, Cys112 and Met121 are responsible for the spectroscopic properties and reduction potential of azurin respectively.[1-6,8-12,18,19]

Site-directed mutagenesis (SDM) of Cys112 to any other proteinogenic amino acid results in species that lack the spectral characteristics associated with blue copper proteins.[20-23] Only a single mutation, Cys112Asp, retains any spectroscopic characteristics.[22] However, instead of wild-type (WT) blue copper type I character, it displays spectroscopic signatures of type 2 copper-binding proteins.[22] Recently the Gray laboratory has developed a "new" copper protein from the C112D azurin mutant.[24] By adding an additional mutation, M121L, a new protein termed Type 0 Cu was created.[24] The variant was capable of binding copper while maintaining a high redox potential and similar geometry to WT azurin all without having a thiolate within the metal binding center.[24] The spectral similarities are attributed to increased interaction of the other ligands to the Cu, especially the Gly45 backbone carbonyl. While the protein displays similar overall geometry and high redox potentials, features not seen in the Cys112Asp mutant alone, the LMCT band is completely absent.[24]

In contrast to Cys112 mutations, mutation of the equatorial His residues or Met^{121} yields proteins dominated by the Cys112 LMCT and retaining blue color.[13,20-22] These SDM studies firmly suggest that Cys112 is essential for blue copper properties. Previous reports from our laboratory using expressed protein ligation (EPL) to incorporate the non-

standard amino acid selenocysteine (Sec) at position 112 have thus far remained the only reported mutations of Cys112 that retain WT type 1 copper site characteristics.[1,3,4] The Cys112Sec variant firmly established EPL as a new and interesting method to probe the type 1 copper site in ways that cannot be accomplished using site-directed mutagenesis.

While the Cys112Sec mutation was able to show electronic contributions of Cys112 to the function of azurin, it was unable to establish any additional information supporting the rack hypothesis.[1,3,4] Thus to investigate how the azurin scaffold reacts to mutations affecting proper site formation, a longer derivative of Cys, homocysteine (Hcy), was incorporated at position 112. Because homocysteine contains a thiolate like Cys, there should be no electronic differences. However, the additional methylene group of homocysteine will push the thiolate closer to the Cu, affecting the geometry and forcing the site to adopt alternative conformations. By analyzing the changes in spectroscopy and reduction potential of the Cys112Hcy azurin variant new insights into the rack hypothesis are possible that cannot be accomplished through standard SDM.[13-15]

2.2 Construction of Cys112Hcy Azurin Using EPL

Ligation with homocysteine at the ligation junction, though retaining a free thiolate for transthioesterification, is not identical to performing EPL with Cys at the ligation junction. The elongated aliphatic chain of homocysteine participates in the ligation, resulting in an intermediate six-membered ring rather than the five-membered ring found in standard ligations using Cys.[25] The different rates of forming six-membered rings compared to the five-membered ring found in Cys ligation could pose difficulties during both steps of EPL. Nonetheless, native chemical ligation (NCL) has been

performed using homocysteine in place of Cys at the ligation point and has shown to proceed smoothly.[25,26]

Though homocysteine was used at the ligation point in NCL, its incorporation into proteins to study protein characteristics has yet to be explored.[26] Using homocysteine at the ligation point in NCL was performed to remove the residual thiolate at the ligation junction by creating Met residues through subsequent methylation of the free thiolate with methyl-iodide.[25] While useful by creating another potential break point other than Cys for NCL, the Cys112Met mutation fails to maintain blue copper properties. Thus no reports incorporating homocysteine to study thiol interactions exist. A single report using homocysteine at the ligation junction (without subsequent methylation) was completed to study the constraining effects of disulfide bonds on cyclic peptides.[25] Consequently, the effects homocysteine may have on metal binding in a metalloprotein has yet to be explored.

Incorporating homocysteine to study metalloprotein geometry could support conclusions of the rack hypothesis that cannot be tested using other techniques. If metalloproteins do use scaffold folding to impose entatic structure to their metal binding sites, an additional methylene unit on Cys112 in azurin should impose difficulties for the protein to correctly adopt the type 1 geometry.[9,15] Thus either the protein will be unable to accommodate the elongated homocysteine side chain and the azurin properties will be perturbed, or the azurin scaffold will adjust to accommodate the elongated side chain and minimize modification to the site's properties.

Using methodology similar to other reported nonproteinogenic amino acid incorporation studies into azurin, a peptide consisting of the last 17 amino acids of azurin

containing a Cys112Hcy mutation (H_2N-HcyTFPGHSALKMGTLTLK-OH) was prepared using standard solid phase peptide synthesis.[1-5] Truncated azurin (residues 1-111) was ligated in-frame to the modified Mxe-intein, and a CBD (NEB – PtxB1 vector) and expressed in an *E. coli* expression system. The fusion protein was then isolated using chitin resin and EPL completed in the presence of mercaptophenyl acetic acid (MPAA) as the transthioesterification mediator to obtain the desired full-length Cys112Hcy azurin variant. The Cys112Hcy Az-apo protein was analyzed using matrix-assisted laser desorpotion ionization mass spectrometry (apo-Cys112Hcy Az: calcd Mw, 13945; obsd Mw, 13948; holo-C112Hcy Az: calcd Mw, 14023; obsd Mw, 14027).

2.3 Spectroscopic Characterization of Cys112Hcy Azurin

Addition of Cu(II) to the apo-Cys112Hcy protein resulted in a blue color change and absorption at 618 nm by visible spectroscopic analysis; a 7 nm blue shifted peak from the WT absorption at 625 nm (Figure 2.1). As previously noted, a strong S(Cys112)3pπ-Cu(II)-$3d_x^2$-$_y^2$ charge transfer results in an intense absorption at 625 nm in WT azurin.[9,27] It is likely that the absorption at 618 nm by the Cys112Hcy Az variant is a blue-shifted S(Hcy112)→Cu(II) charge transfer band. This shift likely represents a perturbation of the site geometry rather than any changes in the Cu(II)-Hcy covalent interaction.

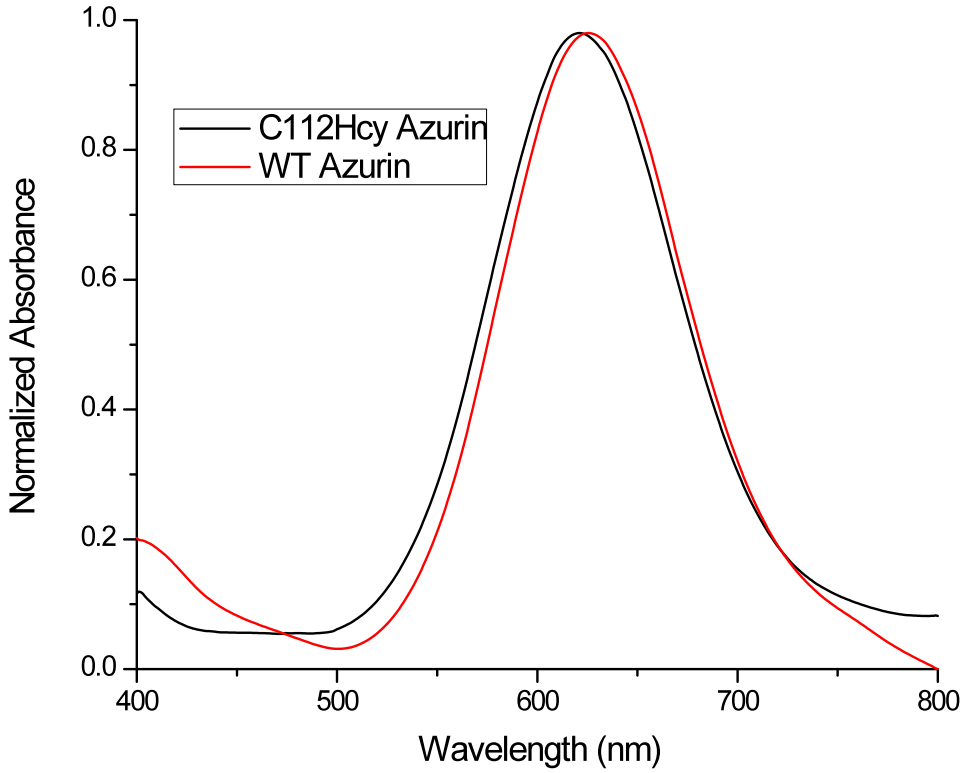


Figure 2.1: Visible spectroscopy of Cys112Hcy azurin: The S_{Hcy112}-Cu(II) charge transfer is at 618 nm. The WT azurin charge transfer is at 625 nm. The peaks are normalized to better display changes in absorbance.

The X-band EPR spectrum (Figure 2.2) of the mutant has a larger parallel hyperfine splitting (58 x 10^{-4} cm^{-1}) than WT azurin (26 x 10^{-4} cm^{-1}) suggesting a reduced covalency between the homocysteine ligand and the Cu(II). Simulation of the spectrum gave *g*-values of 2.24, 2.045, 2.076 (compared with 2.26, 2.056, and 2.039 for WT azurin).[3,4,9,27] The site therefore seems to maintain its overall axial character suggesting little difference between the homocysteine and WT type 1 copper site geometries.

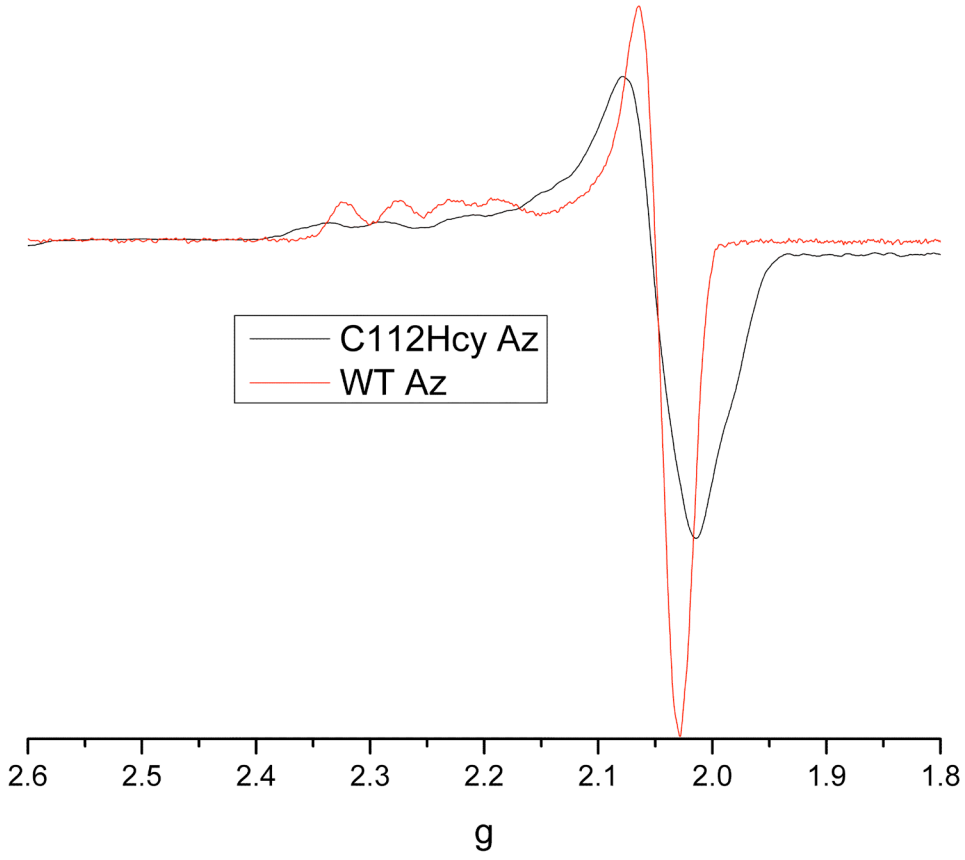


Figure 2.2: X-band EPR spectrum for Cys112Hcy azurin overlaid with WT azurin. EPR Conditions: Field Center = 3000 G, Field Sweep = 2000 G, Modulation Amplification = 3 G, Frequency = 9.049 GHz, Power = 0.2 mW, Temperature = 30 K.

2.4 Electrochemical Analysis

An observed reduction potential increase of 35 mV (360 mV Cys112Hcy, 325 mV WT Az vs NHE) is also suggestive of decreased covalent character between the S(Hcy112)-Cu(II) in correlation with the EPR data presented (Figure 2.3). The increased potential may be a result of geometric compensation and may be allowing more Met-Cu(II) interaction and thus an increased potential.

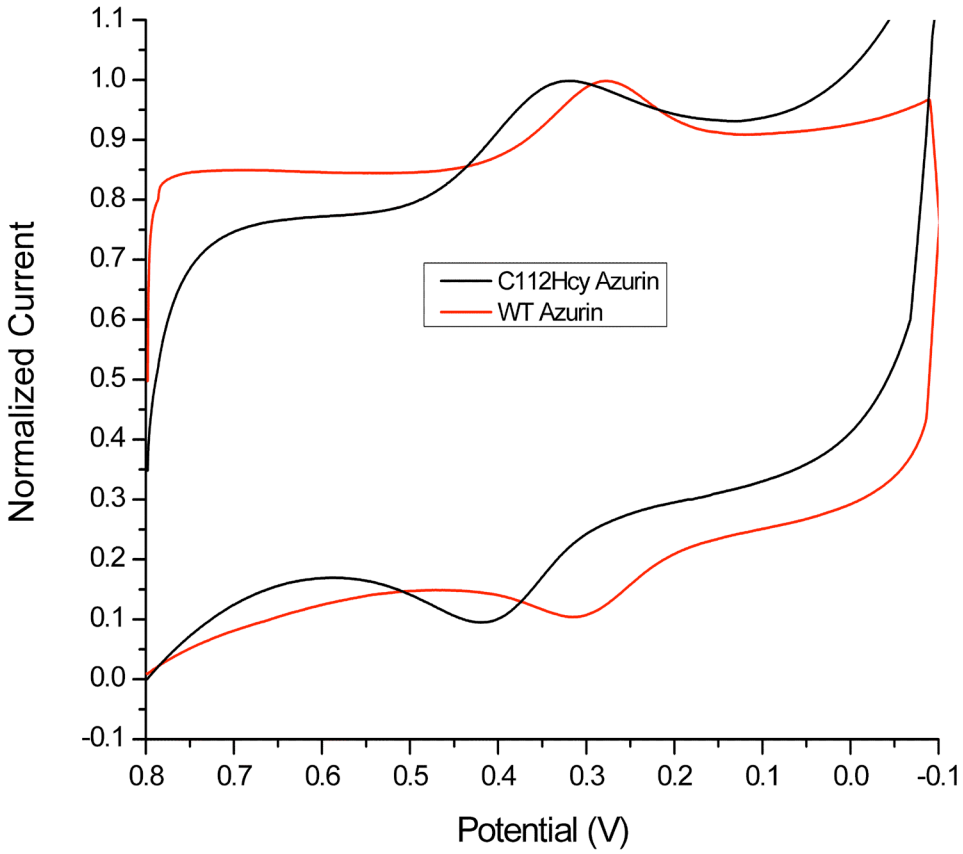
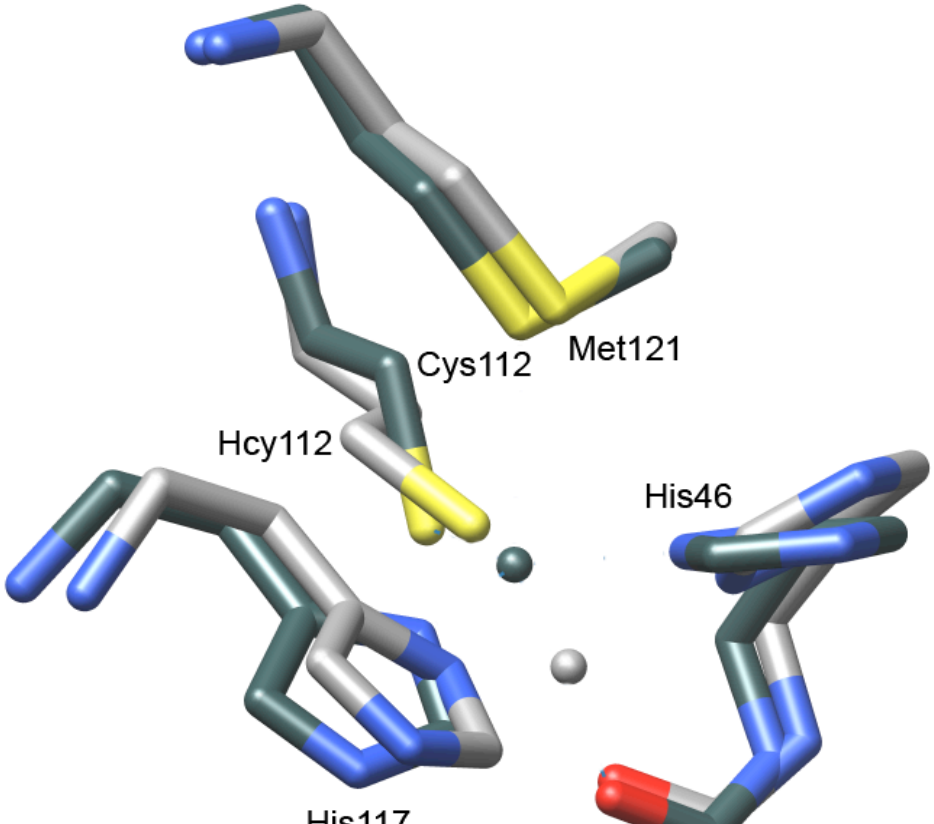


Figure 2.3: Cyclic voltammetry overlay of WT azurin and Cys112Hcy azurin. The potential for the mutant is 35 mV higher than WT under identical conditions (NH_4OAc, pH 5.0, 0.1 M NaCl.). WT Azurin $E^1/_2$ = 325 mV. C112Hcy Az $E^1/_2$ = 360 mV. Values are corrected to NHE.

2.5 Molecular Modeling

To further validate geometric changes in the Cys112Hcy mutant, molecular modeling using VMD software was undertaken with the help of Nicholas Marshall from the Lu group. Figure 2.4 clearly shows that the overall site distortion is minimal when compared to the WT overlay, supporting the claim that the site has undergone geometric rearrangement in order to maintain type 1 copper characteristics. As modeled, the Cu appears to be at a greater distance to Met121 (4.4 Å) than the corresponding interaction in WT azurin (3.18 Å). Though demonstrating very little perturbation of the active site

residues, VMD is currently not capable of fully simulating M^{2+}-ligand charge interactions and thus models solely on electrostatics, altering the actual Met121 and Gly45 distances. As such, it is highly likely that the Cu positioning is very similar in both the variant and WT azurin, explaining the minor 7 nm blue shifted spectra and expanded hyperfine splittings seen in the X-band EPR spectrum.



	Bond Length (Å)	
	WT Azurin	Hcy Azurin
Gly45 - Cu	2.84	2.10
His46 - Cu	1.99	2.12
His117 - Cu	2.11	2.23
Met121 - Cu	3.18	4.40
(Cys)(Hcy)112 - Cu	(Cys)2.27	(Hcy)2.22

Figure 2.4: Molecular model structural overlay of Cys112Hcy azurin with WT azurin. Cys112Hcy is in light gray, WT azurin is in dark gray. The approximated bond lengths to the modeled Cu are listed in the provided table.

2.6 CONCLUSIONS

An azurin variant containing a Cys112Hcy mutation has been obtained through EPL. The mutant displays typical type 1 copper site features but with an apparent decrease in the covalent interaction between the S(Hcy112)-Cu(II) when compared with the WT S(Cys112)-Cu(II) interaction. While not conclusive, incorporation of homocysteine in azurin lends support to the "rack hypothesis." The Cys112Hcy variant maintained its blue color suggesting little perturbation of the site. Additionally, while the hyperfine splittings of the variant were wider than those of WT azurin, suggesting weaker thiolate-Cu interaction, they were still of the same magnitude and maintained similar *g* values and axial geometry. Thus the EPR data further support that little geometric variation is induced in the Cys112Hcy azurin variant, suggesting that the scaffold was able to adapt its structure to maintain WT properties.

Molecular modeling, while limited due to current computer programming capabilities, further supports that the azurin scaffold was capable of rearranging to maintain type 1 geometry and the cupredoxin fold. It appears that the protein compensates for the additional methylene group on homocysteine by pushing the ligand side chain deeper into the protein cavity, and thus maintaining a similar S-Cu(II) distance to WT azurin.

Because determining electron transfer kinetics is a difficult set of experiments to do and requires a substantial amount of protein, the electron transfer kinetics of the variant were not measured. Instead, to gauge any changes to WT properties, the redox behavior of the variant was compared to WT azurin. While the reduction potential is 35 mV higher than tthat of he WT, it is certainly not suggestive of large changes to the site.

2.7 FUTURE DIRECTIONS

Validation of Site Geometry. The spectroscopic studies thus completed while suggestive of little perturbation of site geometry, are not conclusive. Studies are needed to verify site geometry to ensure that the Cys112Hcy variant molecular modeling data is reliable. Verifying the molecular modeling will provide more evidence supporting the rack hypothesis. EPL derived metalloproteins are notoriously difficult to crystallize because of misfolded protein species and the low protein yield following the ligation. While challenging, obtaining a crystal structure would definitively assign bond lengths and angles supporting any changes to the site's geometry. Additionally, our laboratory has previously established collaborative efforts using ENDOR spectroscopy to establish bond lengths. These collaborative efforts could provide additional means to determine perturbations present in the variant scaffold.

Lastly, while difficult to perform, determining the electron transfer kinetics of the site in order to obtain approximate reorganization energy values would provide evidence supporting the rack hypothesis. While all of the current spectroscopic evidence supports that the azurin scaffold successfully accommodated the elongated side chain of homocysteine, without verification of electron transfer kinetics the true changes to the site will remain unknown. Key to the rack hypothesis is that the reorganization energy of blue copper proteins remains relatively static because of the entatic forces imposed by the scaffold. Only kinetic studies can provide that information.

2.8 EXPERIMENTAL

General. All chemicals used were obtained from Aldrich. All amino acids were purchased from Novabiochem, Advanced Chemtech, or Chem-Impex Internation. Protecting groups used for Fmoc-SPPS were Cys(Trt), Ser(OtBu), Glu(OtBu), Cys(StBu), Asn(Trt), His(Trt), Lys(Boc), Thr(OtBu). Methylene chloride (DCM) was distilled from calcium hydride. Tetrahydrofuran (THF) was distilled from sodium and benzophenone. Dimethylformamide (DMF) was dried over 2 Å sieves before use. All aqueous buffers were degassed under Ar or N_2 as specified immediately prior to use. UV-Visible, MS, and EPR Spectroscopy: UV-Vis spectra were taken on a Cary 5000 spectrometer. X-band EPR spectra were collected on a Varian E-122 spectrometer at the Illinois EPR Reseach Center (IERC). The samples were run at ~30 K using liquid He and an Air Products Helitran cryostat without glassing agents. The magnetic field was calibrated using a Metrolab PT 2025 NMR Gaussmeter, and the microwave frequency was measured with an EIP model 578 frequency counter equipped with a high-frequency option.

Mass spectral data was collected at the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois by either Electrospray Ionization (ESI MS), or matrix assisted laser desorption ionization (MALDI) techniques as indicated. Mass spectral data reported as – mass, peak (percent).

Preparative RP-HPLC was done with a Waters Delta 600 system. Solution A was 0.1 % TFA in H_2O, and solution B was 80 % MeCN/20 % H_2O with 0.1 % TFA. Unless otherwise stated, a linear gradient of 20% to 70% B over 23 min was used for all runs.

Solid phase peptide synthesis (SPPS) was performed on a Rainin PS3 automated peptide synthesizer or manually using medium fritted funnels.

H_2N-HcyTFPHGSALMKGTLTLK-COOH

Synthesis of C112Hcy 17-mer Peptide, (3.1), Lab Book 7, page 21.

Solid Phase Peptide Synthesis. The peptide was synthesized on a 0.2 mmol scale with a Rainin model PS3 peptide synthesizer (Woburn, MA) using standard Fmoc-based chemistry. Wang resin preloaded with Fmoc-Lys(Boc) was pre-swelled in DMF (6 x 10 min, 6 mL). All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Synthesis of the peptide from Lys128 through Thr113 was completed using a five-fold excess of amino acids pre-activated in 0.4 M N-methylmorpholine (NMM) (6 mL) and O-(1H-6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HCTU) (4 equiv.) for 3 min and coupled to the resin for 45 min under N_2 sparging. Amino acids were double coupled when necessary and completion of the couplings were monitored by qualitative Kaiser test. Fmoc-Hcy(Trt)OH (3 equiv.) was coupled without pre-activation using 1-hydroxybenzotriazole hydrate (HOBt) (6 equiv.) and 1-ethyl-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI) (6 equiv.) in DMF at 25 °C for 2-4 h until full coupling could be observed by qualitative Kaiser tests.

Cleavage. The resin carrying the fully protected peptide was washed with ethanol (3 x 15 mL) and dried under reduced pressure. Peptide cleavage from the resin was achieved by treating the resin with a cocktail of 5% triisopropylsilane (TIPS)/H_2O/thioanisole in neat trifluoroacetic acid (TFA) (10 mL) at 25 °C for 2 h. The

mixture was filtered and the eluent concentrated using rotary evaporation to obtain a yellow oil. The peptide was precipitated from the yellow oil by trituration with cold diethyl ether and isolated by centrifugation at 3000 x g at 10 ˚C. The crude peptide was dissolved in 1:1 dH_2O/acetonitrile (MeCN) (~10-15 mL) and lyophilized prior to purification. The crude lyophilized product was purified by preparative RP-HPLC and lyophilized directly after purification to afford a white solid. (Yield: 65 mg). MALDI-MS. Expected: 1819.14, Observed: 1819.5.

Expressed Protein Ligation, Lab Book 7, page 48.

Cultures of *E. coli* BL21 (DE3) cells containing the NEB pTXB1 plasmid to express the azurin(1-111)-Intein-CBD fusion protein were grown in LB media for 8 h at 37 ºC and used to inoculate eight 2 L flasks of LB media containing 100 mg/mL ampicillin. The cells were grown at 37 ºC for 16 h with shaking at 210 rpm and harvested at 9000 x g. The cell stock was then frozen stored at -80 ºC until needed.

The frozen cell stock was re-suspended in a lysis buffer containing 20 mM HEPES, pH 8.0, 250 mM NaCl, 1 mM EDTA, 1 mM PMSF, 0.1% Triton-X-100 and 1 M urea (100 mL). The suspension was then lysed using sonication (Misonix Sonicator 4000, 0.5 inch diameter probe) for a work time of 6 min (6 sec. on, 10 sec. rest). Additional urea was added to bring the final concentration to 1.5 M and the lysate was subjected to a second round of sonication for 3 min work time (6 on, 10 off). The crude lysate was centrifuged at 13,500 x g for 30 min. The supernatant was carefully decanted and the fusion protein bound by batch absorption to 70 mL of chitin resin pre-equilibrated with 20 mM HEPES, pH 7.2, 250 mM NaCl, and 1 mM EDTA (buffer 1) for 1-2 h at 4 º C.

The chitin resin was then poured into a column and the column headspace was purged with Ar. The column was then washed with 5 column volumes of buffer 1 by cannulation under Ar pressure.

Ligations were initiated by the addition of the 17-mer peptide (1.02 mM, 60 mg) and 2 equiv tris-(2-carboxyethyl)phosphine (TCEP) (3 mM, 30 mg) in 40 mL of degassed buffer 1 (35 mL) containing ~50 mM mercaptophenyl acetic acid (MPAA) (50 mM, 350 mg) to the column under Ar pressure via cannulation. The chitin resin was re-suspended in the column and the entire column was agitated gently at 4 ºC for 64 h. After ligation, the column was eluted under Ar pressure and washed with 1 column volume of buffer 1. The eluent was centrifuged at 13500 x g for 30 min and the supernatant was concentrated using 10,000 MWCO Centricon concentration spin tubes at 3000 x g to a final volume of ~10 mL. The concentrated protein was then exchanged into 50 mM Universal buffer (50 mM NaOAc, 40 mM each MES/MOPS/TRIS/CAPS) pH 7 via PD10 (G25 Sephadex) columns. The desalted sample was titrated with incremental additions of freshly made 20 mM $CuSO_4$ solution in H_2O. The titrated sample was purified using a Pharmacia SP HiTrap column (5 mL) that was washed with 10 column volumes of 1 M NaCl in H_2O and equilibrated with 10 column volumes of 25 mM MOPS, pH 6.5. Following purification, the protein was concentrated and analyzed using MALDI-TOF-MS and ESI-MS. (Apo-calculated: 13930. Observed: 13932. Holo-calculated: 13994. Observed: 13996).

Electrochemical Measurements, Lab Book 6, page 44.

The reduction potential of each mutant was determined by cyclic voltammetry after verifying the WT azurin reduction potential using a CH Instruments 617A potentiostat equipped with a picoamp booster and a Faraday cage. A pyrolytic graphite edge (PGE) electrode was polished and 2-3 mL protein solution was applied directly to the electrode following previously described methods. After a short incubation time, the electrode was immersed in either 20 mM NaOAc, pH 4.0 with 100 mM NaCl, 20 mM NH_4OAc, pH 5.0 with 100 mM NaCl, or 25 mM KH_2PO_3, 100 mM KCl before data collection. Each protein was then sampled at varying scan rates between 10 mV and 500 mV. The reduction potentials were measured against Ag/AgCl and converted to NHE.

2.9 REFERENCES

(1) Berry, S. M.; Gieselman, M. D.; Nilges, M. J.; van der Donk, W. A.; Lu, Y. "An Engineered Azurin Variant Containing a Selenocysteine Copper Ligand." *J. Am. Chem. Soc.,* **2002**, *124*, 2084-2085.

(2) Berry, S. M.; Ralle, M.; Low, D. W.; Blackburn, N. J.; Lu, Y. "Probing the Role of Axial Methionine in the Blue Copper Center of Azurin with Unnatural Amino Acids." *J. Am. Chem. Soc.,* **2003**, *125*, 8760-8768.

(3) Ralle, M.; Berry, S. M.; Nilges, M. J.; Gieselman, M. D.; van der Donk, W. A.; Lu, Y.; Blackburn, N. J. "The Selenocysteine-Substituted Blue Copper Center: Spectroscopic Investigations of Cys112SeCys *Pseudomonas aeruginosa* Azurin." *J. Am. Chem. Soc.,* **2004**, *126*, 7244-7256.

(4) Sarangi, R.; Gorelsky, S. I.; Basumallick, L.; Hwang, H. J.; Pratt, R. C.; Stack, T. D. P.; Lu, Y.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. "Spectroscopic and Density Functional Theory Studies of the Blue-Copper Site in M121SeM and C112SeC Azurin: Cu-Se Versus Cu-S Bonding." *J. Am. Chem. Soc.,* **2008**, *130*, 3866-3877.

(5) Garner, D. K.; Vaughan, M. D.; Hwang, H. J.; Savelieff, M. G.; Berry, S. M.; Honek, J. F.; Lu, Y. "Reduction Potential Tuning of the Blue Copper Center in *Pseudomonas aeruginosa* Azurin by the Axial Methionine as Probed by Unnatural Amino Acids." *J. Am. Chem. Soc.,* **2006**, *128*, 15608-15617.

(6) Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z.; Guo, Z.; Lu, Y. "A Site-Selective Dual Anchoring Strategy for Artificial Metalloprotein Design." *J. Am. Chem. Soc.,* **2004**, *126*, 10812-10813.

(7) Marshall, N. M.; Garner, D. K.; Wilson, T. D.; Gao, Y.-G.; Robinson, H.; Nilges, M. J.; Lu, Y. "Rationally tuning the reduction potential of a single cupredoxin beyond the natural range." *Nature,* **2009**, *462*, 113-116.

(8) Gray, H. B.; Malmström, B. G.; Williams, R. J. P. "Copper Coordination in Blue Proteins." *J. Inorg. Biochem.,* **2000**, *5*, 551-559.

(9) Malmström, B. G. "Rack-Induced Bonding In Blue-Copper Proteins." *Eur. J. Biochem.,* **1994**, *223*, 711-718.

(10) Solomon, E. I.; Szilagyi, R. K.; DeBeer George, S.; Basumallick, L. "Electronic Structures of Metal Sites in Proteins and Models; Contributions to Function in Blue Copper Proteins." *Chem. Rev.,* **2004**, *104*, 419-458.

(11) Vallee, B. L.; Williams, R. J. P. "Metalloenzymes: The Entatic Nature of Their Active Sites." *Proc. Natl. Acad. Sci. USA,* **1967**, *59*, 498-505.

(12) Williams, R. J. P. "Energised (entatic) States of Groups and of Secondary Structures in Proteins and Metalloproteins." *Eur. J. Biochem.,* **1995**, *234*, 363-381.

(13) DeBeer, S.; Wittung-Stafshede, P.; Leckner, J.; Karlsson, G.; Winkler, J. R.; Gray, H. B.; Malmstrom, B. G.; Solomon, E. I.; Hedman, B.; Hodgson, K. O. "X-ray absorption spectroscopy of folded and unfolded copper(I) azurin." *Inorg. Chim. Acta.,* **2000**, *297*, 278-282.

(14) Di Bilio, A. J.; Hill, M. G.; Bonander, N.; Karlsson, B. G.; Villahermosa, R. M.; Malmstrom, B. G.; Winkler, J. R.; Gray, H. B. "Reorganization Energy of Blue Copper: Effects of Temperature and Driving Force on the Rates of Electron Transfer in Ruthenium- and Osmium-Modified Azurins." *J. Am. Chem. Soc.,* **1997**, *119*, 9921-9922.

(15) Winkler, J. R.; Wittung-Stafshede, P.; Leckner, J.; Malmstrom, B. G.; Gray, H. B. "Effects of Folding on Metalloprotein Active Sites." *Proc. Natl. Acad. Sci. USA,* **1997**, *94*, 4246-4249.

(16) Karlsson, B. G.; Aasa, R.; Malmstrom, B. G.; Lundberg, L. G. "Rack-induced bonding in blue copper proteins: Spectroscopic properties and reduction potential of the azurin mutant Met121Leu." *FEBS Lett.,* **1989**, *253*, 99-102.

(17) Messerschmidt, A.; Prade, L.; Kroes, S. J.; Sanders-Loehr, J.; Huber, R.; Canters, G. W. "Rack-Induced Metal Binding Vs. Flexibility - Met121His Azurin Crystal Structures At Different pH." *Proc. Natl. Acad. Sci. USA.,* **1998**, *95*, 3443-3448.

(18) Guckert, J. A.; Lowery, M. D.; Solomon , E. I. "Electronic Structure of the Reduced Blue Copper Active Site: Contributions to Reduction Potentials and Geometry." *J. Am. Chem. Soc.,* **1995**, *117*, 2817-2844.

(19) Olsson, M. H. M.; Ryde, U. "The influence of axial ligands on the reduction potential of blue copper proteins." *J. Biol. Inorg. Chem.,* **1999**, *4*, 654-663.

(20) DeBeer, S.; Kiser, C. N.; Mines, G. A.; Richards, J. H.; Gray, H. B.; Solomon, E. I.; Hedman, B.; Hodgson, K. O. "X-ray Absorption Spectra of the Oxidized and Reduced Forms of C112D Azurin from Pseudomonas aeruginosa." *Inorg. Chem.,* **1999**, *38*, 433-438.

(21) Faham, S.; Mizoguchi, T. J.; Adman, E. T.; Gray, H. B.; Richards, J. H.; Rees, D. C. "Role of the Active-site Cysteine of Pseudomonas aeruginosa azurin. Crystal structure analysis of the Cu(II)(Cys112Asp) protein." *J. Inorg. Biol. Chem.,* **1997**, *2*, 464-469.

(22) Mizoguchi, T. J.; Di Bilio, A. J.; Gray, H. B.; Richards, J. H. "Blue to Type 2 Binding. Copper(II) and Cobalt(II) Derivatives Of a Cys112Asp Mutant Of *Pseudomonas aeruginosa* Azurin." *J. Am. Chem. Soc.,* **1992**, *114*, 10076-10078.

(23) Piccioli, M.; Luchinat, C.; Mizoguchi, T. J.; Ramirez, B. E.; Gray, H. B.; Richards, J. H. "Paramagnetic NMR spectroscopy and coordination structure of cobalt(II) Cys112Asp azurin." *Inorg. Chem.,* **1995**, *34*, 737-742.

(24) Lancaster, K. M.; George, S. D.; Yokoyama, K.; Richards, J. H.; Gray, H. B. "Type-zero copper proteins." *Nat. Chem.,* **2009**, *1*, 711-715.

(25) Pachamuthu, K.; Schmidt, R. "Synthesis of Methionine Containing Peptides Related to Native Chemical Ligation." *Synlett,* **2003**, 659-662.

(26) Chatterjee, C.; Patton, G. C.; Cooper, L.; Paul, M.; van†der†Donk, W. A. "Engineering Dehydro Amino Acids and Thioethers into Peptides Using Lacticin 481 Synthetase." *Chem. Biol.,* **2006**, *13*, 1109-1117.

(27) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. "Electronic Structures of Active Sites in Copper Proteins: Contributions to Reactivity." *Chem. Rev.,* **1992**, *92*, 521-542.

CHAPTER 3

INCORPORATING ^{77}SE-SEC INTO AZURIN: C112U AZURIN[i]

3.1 INTRODUCTION

Protein design and engineering has contributed greatly to our understanding of protein structure and function. A common practice when designing and engineering new function into protein scaffolds is to use site-directed mutagenesis (SDM) to replace a specific amino acid in proteins with other natural amino acids. While quite powerful, it has become increasingly evident that SDM cannot address the precise role of a specific amino acid because, with limitation of 20 natural amino acids, it is often difficult to change only a single factor. Thus when a residue is altered, both the electronics and sterics of the site are perturbed, limiting the ability to precisely determine function. This limitation is even more evident in metalloprotein design and engineering, as even a smaller number of natural amino acids are capable of coordinating to metal ions.[1,2]

More importantly, metal-binding sites are intricately designed to accept a specific metal ion among large numbers of other metal ions for binding and activity. Since geometry and ligand donor sets of metal-binding sites are much more varied than those of metal-free active sites, a subtle change of one factor can have a profound effect on structure and activity. Therefore the need for introducing isostructural unnatural amino acids into metalloproteins is even greater. Such an endeavor not only defines the role of

[i] This work was done in collaboration with Dr. David Britt with help from Michelle Dicus, PhD at the University of California-Davis. All ENDOR data, figures, and analysis were generated by Michelle Dicus, PhD.

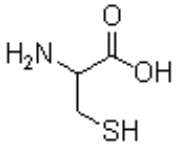
amino acids more precisely, but also produces novel metalloproteins with both predicted and unprecedented properties.[3,4]

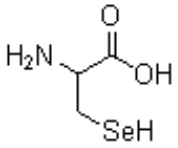
The type I site is defined by an unusually short copper to cysteine bond (2.1 Å) and this shortened bond, together with its unique metal-binding site geometry, is responsible for an intense ligand to metal charge transfer (LMCT) band resulting from the S^{γ}(Cys112) 3pπ → Cu(II) d_x^2-$_y^2$ charge transfer transition giving azurin its strong visible absorption at 625 nm and thus its intense blue color.[5-7] The highly covalent nature of the LMCT between the copper and Cys ligand is also responsible for the narrow hyperfine splittings in the parallel region of the EPR spectra ($A_{\parallel}$ ~ 60 X 10^{-4} cm^{-1}), smaller than those of "normal" copper coordination compounds (($A_{\parallel}$ ~ 160 X 10^{-4} cm^{-1}) (See Figures 1.4, 1.5).[7]

3.2 INCORPORATING SEC INTO AZURIN: BACKGROUND

Precisely studying the role of a single residue requires incorporating nonproteinogenic amino acid analogues. Selenocysteine (Sec or U), an isosteric analogue of Cys, has been used to study the type I site of azurin.[8-10] Since Se and S share similar atomic radii, electronegativity, and oxidation states, the replacement by Sec is much more conserved than replacement of Cys with other naturally occurring amino acids (Figure 3.1). The similarity also yields comparable oxidized analogues (e.g. disulfides and diselenides) and allows for mixed S-Se compounds,[11,12] a feature that the natural amino acid Ser cannot mimic.[13] However, although both Se and S are Group VIA elements with similar physical characteristics, they exhibit different reactivities. Most notably, the pK_a

of Sec (5.2) is much lower than that of Cys (8.6), suggesting that at physiological pH Sec will be in its anionic selenolate form rendering it more nucleophilic than Cys.[11,14] Unfortunately the replacement of Cys with Sec is not easily accomplished by site directed mutagenesis and requires the use of alternative techniques for its incorporation (see Chapter 1, section 1.6).





	Cysteine *Sulfur*	**Selenocysteine** *Selenium*
Electron Configuration	[Ne] $3s^2.3p^4$	[Ar] $3d^{10}.4s^2.4p^4$
Atomic Radius (pm)	100	115
Covalent Radius (pm)	102	116
Bond length of dichalcogen (pm)	205 (S-S)	232 (Se-Se)
Electronegativity	2.58	2.55
Redox Potential at pH <7 (mV)	231	276
Common Oxidation States	-2, 0, +4, +6	-2, 0, +4, +6
pKa	8.3	5.2

Figure 3.1: Comparison of the physiochemical properties of cysteine and selenocysteine.

Using EPL Sec was successfully incorporated into azurin in place of Cys112 (C112U Azurin).[8-10] The C112U azurin variant demonstrated distinct differences from the WT protein.[8-10] The first characteristic change was a red shift in the UV-Vis spectrum from 625 nm for WT to 677 nm for the C112U mutant likely resulting from the lower ionization energy of Se when compared to S.[8,9] More dramatically are the changes noted in the EPR spectra between the two proteins.[9] A marked increase of the parallel hyperfine splitting (100 G) of the C112U mutant compared to WT azurin (56 G) suggests that the Se-Cu bond is less covalent in nature than the S-Cu bond.[9] Furthermore, the observed

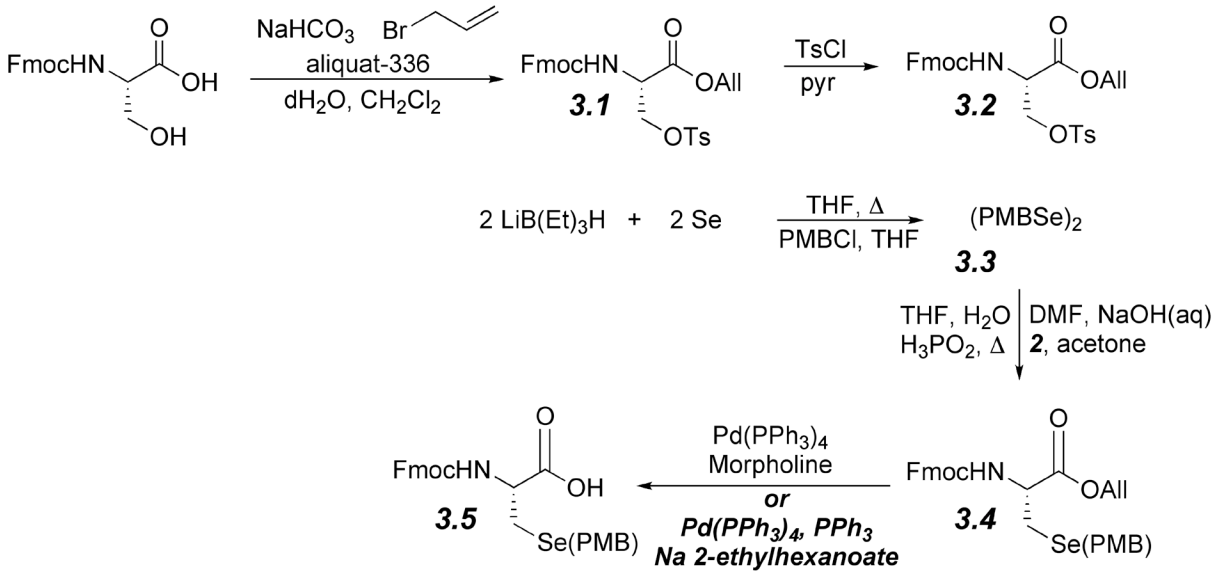
increased rhombic character of the site with a larger increase in g_y than in the g_z tensor is consistent with increased axial interaction or loss of Cu(II) interaction with the selenolate.[9] This finding was expected as Se is softer than S and thus would prefer coordination with Cu(I) rather than Cu(II). Surprisingly however, no dramatic change was noted in the reduction potential of the site when the ligand was changed from Cys to Sec.[8]

Though the C112U azurin mutant provided a unique view of the type I site, many questions still stand regarding the electron densities residing on the Cu and ligands in the azurin type I site.[15] Thus it was proposed that incorporating an isotopically enriched ^{77}Se-Sec analogue into azurin could provide a means through Electron-Nuclear Double Resonance (ENDOR) spectroscopy to empirically evaluate these densities. Such a ^{77}Se-Sec derivative is prohibitively expensive to synthesize. Therefore, with the low yields of protein obtained during the semi-synthesis of non-enriched C112U azurin, it became evident that optimization of the synthetic and EPL schemes would be necessary to justify incorporation of an expensive ^{77}Se derivative.

3.3 Optimization of the Synthesis of Sec(PMB).

In order to accomplish substitution of Cys112 with selenocysteine, a derivative of selenocysteine was needed that was suitable for standard Fmoc-based Solid Phase Peptide Synthesis (SPPS). The most common derivative of selenocysteine used for SPPS has been Fmoc-Sec(PMB)-OH (PMB: *p*-methoxybenzyl) primarily due to incompatibilities with the trityl (Trt) protecting group.[16-18] Many procedures have been

developed over the years to synthesize Fmoc-Sec(PMB)-OH from various starting materials. In recent years, a scheme originating from the van der Donk laboratory has emerged as a convenient methodology for Sec(PMB) synthesis.[19] While this route is already highly optimized (Scheme 3.1), the final allyl deprotection step suffers from incidental Fmoc-deprotection when using morpholine as the allyl transfer agent. By changing the allyl transfer reagent to sodium 2-ethylhexanoate, the partial removal of the Fmoc group was eliminated and yields were modestly increased (99% from 95%).



Scheme 3.1: Synthetic route for Fmoc-Sec(PMB)-OH (***3.5***) from Fmoc-Ser-OH.[19] Conditions in **bold** are modified and optimized conditions used for allyl deprotection differing from ref 19.

Starting from the inexpensive Fmoc-Ser-OH starting material, the precursor Fmoc-Se(PMB)-Sec-OAll was prepared in 3 steps (Scheme 3.1). Selective installation of the allyl ester was followed by activation of the alcohol with *p*-toluenesulfonyl chloride. Installation of the *p*-methoxylbenzyl selenol was accomplished through reduction of the

$(PMBSe)_2$ and displacement of the *p*-tolunenesulfonate group in basic conditions. The final optimized step to selectively remove the allyl ester protection was carried out as described by McCombie with slight modifications and by using sodium 2-ethylhexanoate as the allyl scavenger rather than morpholine to reduce undesired Fmoc-deprotection (see experimental).[20] After successful synthesis of Fmoc-Sec(PMB)-OH, the final amino acid derivative was incorporated into the C-terminal 17 amino acid sequence of azurin using standard Fmoc-based chemistry.

3.4 Solid Phase Peptide Synthesis of C112U 17-mer

Substitution of Cys112 with selenocysteine requires a synthetic peptide with an N-terminal Sec residue that participates directly in the ligation reaction, substituting for the role of cysteine. The use of Sec residues to facilitate the trans-selenoesterification needed for peptide ligation in EPL has been previously been reported. Hondal showed that replacement of Cys110 in RNase A with Sec could be accomplished using a synthetic 15 amino acid peptide containing an N-terminal Sec residue.[21] The EPL product protein was indistinguishable from the WT protein with similar rates of catalysis indicating that replacement of Cys with Sec does not dramatically perturb either function or protein folding. More recently, Hondal has demonstrated Sec incorporation into High M_r thioredoxins without disruption of protein folding.[22,23] Concurrently, the van der Donk lab developed and used similar methods for azurin.[8,9,19]

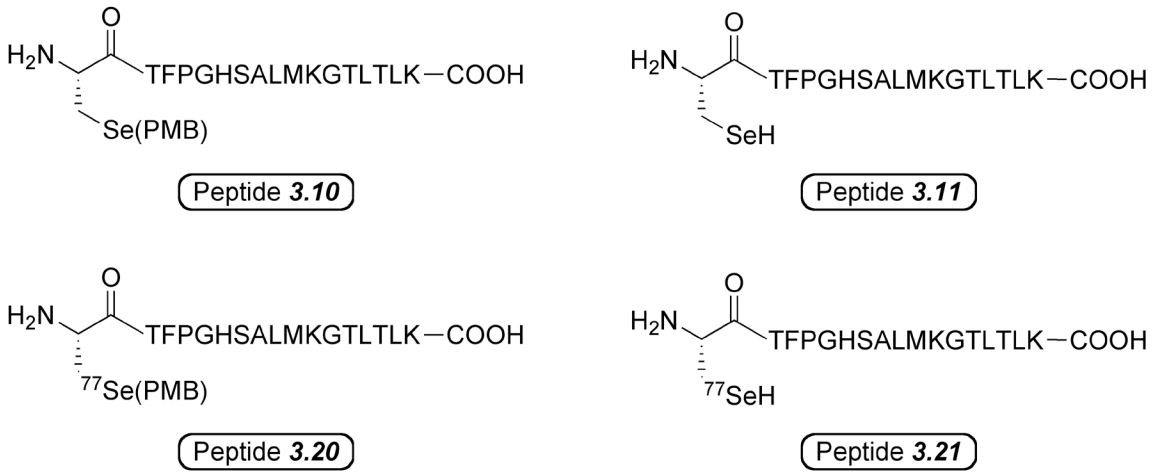
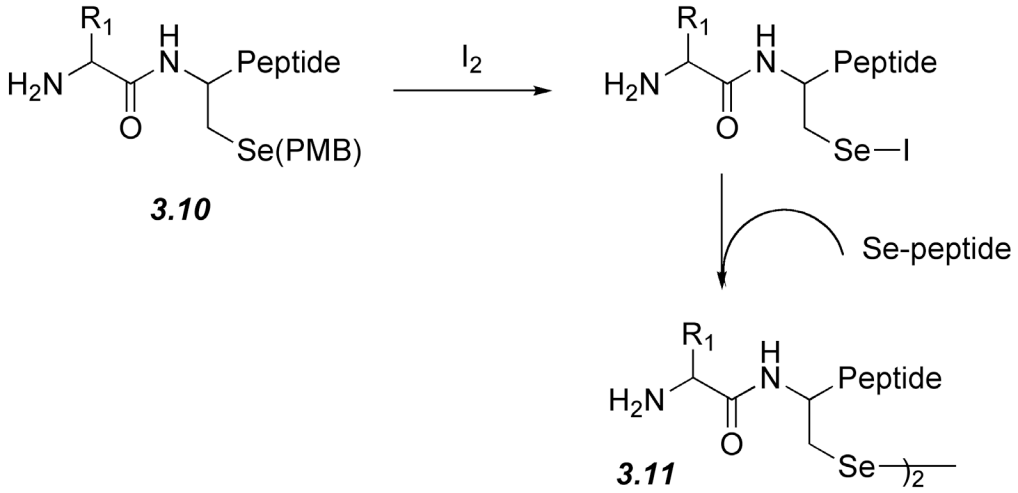


Figure 3.2: Azurin 17-mer peptides for incorporating selenocysteine at position 112 (C112U Az).

The C-terminal 17-mer of azurin was synthesized as reported using Fmoc-SPPS but with a cysteine to selenocysteine mutation at position 112 (H_2N-$U_{(PMB)}$TFPGHSALMKGTLTLK-OH) to make peptide ***3.10*** (Figure 3.2).[8] Because the PMB protecting group is resistant to removal using standard peptide cleavage cocktails like H_2O/triisopropylsilane(TIPS)/ethanedithiol (EDT) in TFA, an alternative procedure was used to afford the free selenol needed for the ligation. Gieselman reported that oxidative cleavage of the PMB group could be accomplished with I_2 in AcOH and was the preferred method for PMB deprotection in early C112U azurin studies.[19] However, this procedure is known to cause deselenization and was noted to decrease overall peptide yields. Additionally, other reagents have been reported to successfully deprotect PMB protected selenols and thiols but many of these reagents are harsh, do not result in full deprotection of the thiol/selenol, and promote deselenization.[24-26] A review of these

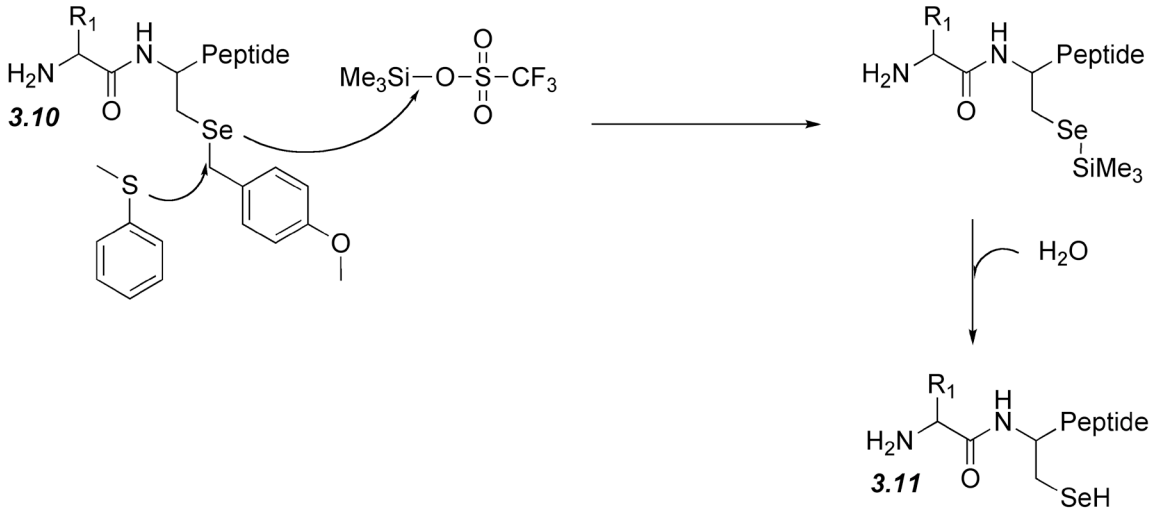
techniques was published explaining the advantages and disadvantages of each of these techniques.[27,28]

The first method to remove the PMB group, affording peptide ***3.11***, was following the deprotection scheme from Gieselman by oxidative cleavage using I_2 in AcOH (Scheme 3.2). The peptide was dissolved in 50% AcOH in 1:1 ACN/H_2O containing I_2 and stirred at room temperature (RT) under N_2 for 5 min. The peptide was then precipitated with cold Et_2O and purified via RP-HPLC. This method yielded the selenol monomer (m/z = 1852.5) as well as the diselenide dimer (m/z = 3703) but contained significant levels of deselenated product (m/z = 1771). Peptide ***3.11*** is very susceptible to deselenization using these conditions as the Se-C bond is weaker compared to the selenyliodide anion[27,28] (Scheme 3.2). Using I_2 in AcOH yielded product that while PMB deprotected, resulted in high amounts of deselenated peptide and thus decreased overall yield of the desired C112U 17-mer peptide.



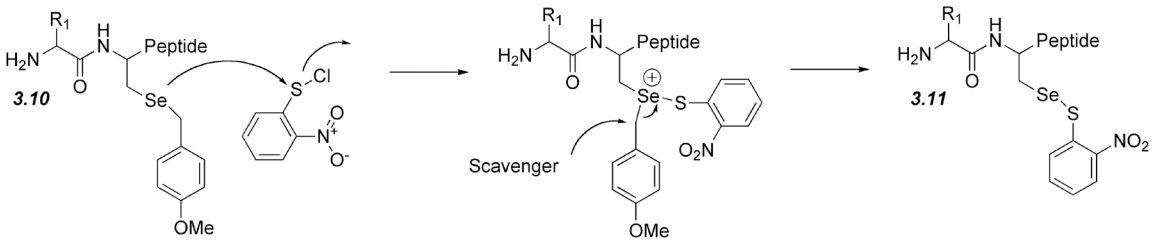
Scheme 3.2: Oxidative PMB deprotection of peptide ***3.10*** with I_2 in AcOH to synthesize peptide ***3.11***.

A second method to deprotect the PMB group is the use of catalysis with strong Lewis acids such as trimethylsilyl trifluoromethanesulfonate (TMSOTf) (Scheme 3.3). The unpurified PMB protected peptide ***3.10*** was dissolved in a ratio of 50:120:690:194 m-cresol/thioanisole/TFA/TMSOTf and stirred under N_2 for 30 min. The TFA and TMSOTf were removed carefully under reduced pressure, the peptide was lyophilized, and purified via RP-HPLC. MALDI-MS analysis showed the desired product (m/z = 1852.06) but also large amounts of uncharacterized impurities. Deselenization occurs rapidly with this method because TMSOTf is a very strong Lewis acid. Furthermore, the reaction product oils out since *m*-cresol is retained with the peptide after extraction with ether. Like the oxidative deprotection with I_2, deprotection with strong Lewis acids results in poor yields and impure material.



Scheme 3.3: Lewis-acid catalyzed removal of PMB protection from peptide ***3.10*** to synthesize ***3.11***.

Removing the PMB protection through formation of a selenyl-sulfide by treatment with 2-nitrophenyl sulfenyl chloride (NpsCl) is an alternative route to harsh chemicals that result in deselenization (Scheme 3.4). However, Se-S bonds can be quite strong and difficult to break, even with reducing agents such as tris(2-carboxyethyl)phosphine (TCEP). However, Hondal and coworkers noted that TCEP can reduce Se-S bonds when used in greater than a two-fold excess.[21,28] Peptide ***3.10*** was treated with NpsCl in 50% MeCN, H_2O, AcOH, and HCl (5:16:1) and stirred at RT for 30 min. (Scheme 3.4). The volatile organics were removed under reduced pressure and the peptide was precipitated and purified via RP-HPLC. Using NpsCl for PMB deprotection results in peptide mixtures consisting of the free selenol peptide ***3.11*** (m/z = 1852), the dimer peptide (m/z = 3704), and the Se(Nps) peptide (m/z = 2006). MALDI MS analysis showed a 1:2 ratio of the SeH:Se(Nps) monomers. Though yielding a ratio of products when using this method, all products are useful in EPL reactions because the Nps protection and diselenide peptide dimers can be reduced with reducing agents such as TCEP in the ligation buffer, freeing the selenol for participation in trans-selenoesterification. Although resulting in functional peptides for EPL reactions, free Nps is very reactive and is known to produce side reactions with Trp residues within peptide sequences.[29] Thus Trp oxidation was noticeable on the EPL ligated product and although such oxidized proteins could be removed during purification, such a deprotection scheme is not optimal.

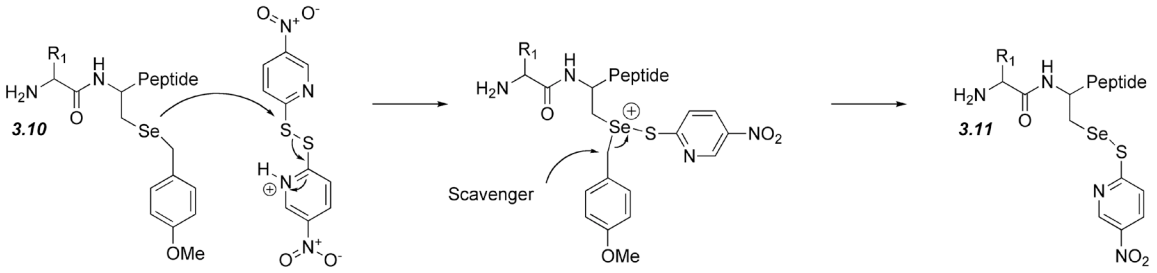


Scheme 3.4: NpsCl mediated removal of PMB protection from peptide *3.10*.

Hondal and coworkers recently developed an alternative selenyl-sulfide exchange group to remove the PMB group from selenols.[27] Instead of NpsCl, 2,2'dithiobis(5-nitropyridine) (DTNP) was used to facilitate PMB removal (Scheme 3.5). Therefore, Wang-resin bound peptide ***3.10*** was treated with 2% H_2O, 2% triisopropylsilane, 2.5% thioanisole, and 1.3 equiv. DTNP in neat TFA at RT for 2 h. Similar to the NpsCl reaction, the DTNP reaction results in peptide mixtures of peptide ***3.11***, the free selenol peptide (m/z = 1852), the dimer peptide (m/z = 3704), and the Se(DTNP) peptide (m/z = 2006). By minimizing the amount of DTNP used, higher yields of the free selenol peptide are achievable. However, an insufficient quantity of DTNP results in incomplete PMB deprotection. The quantity used by Hondal (1.3 equiv.) was sufficient to remove all PMB protection.[27] Using DTNP to deprotect the selenol resulted in a product ratio of 1:3:3:4 deselenization product:free selenol:dimer:Se(pys). All the peptides but the deselenated material are functional in EPL reactions because the Se bond and diselenide dimers are reduced by the TCEP present in the ligation buffers.

Using DTNP to remove the PMB group is a superior methodology to using NpsCl. First, using DTNP yields more free selenol than when using NpsCl. Secondly, in contrast to NpsCl, the released pys group is non-reactive with other residues in the

protein. Finally, unlike NpsCl, where an additional reaction is required to remove the PMB group after cleavage from the resin, DTNP is added during the cleavage from the resin, eliminating the need for a separate reaction. Moreover, NpsCl reactions were usually completed after full purification of peptide ***3.10***, therefore requiring an additional HPLC purification step. With DTNP, only a single HPLC purification step is necessary, saving time and money.



Scheme 3.5: DTNP mediated removal of PMB protection from peptide *3.10*.

Once the purified peptide was available it was assessed by MALDI MS to ensure that Se was present. Because deselenization occurs in the presence of piperidine during the final Fmoc cleavage, verification that the correct mass was attributable to Se was necessary. The C112U 17-mer peptide isotopic distribution ($C_{80}H_{133}N_{21}O_{22}SSe$, Exact Mass: 1851.88) was simulated using the Applied Biosystems Data Explorer program (version 4). The peptide was then checked using MALDI MS TOF analysis in reflective mode to get the monoisotopic mass of the peptide, verifying the correct isotopic distribution (Figure 3.3). The analysis showed that the monoisotopic mass and isotopic distribution of the C112U 17-mer matched the simulated data, thereby verifying the

correct isotopic distribution pattern expected for a peptide containing one selenium atom.

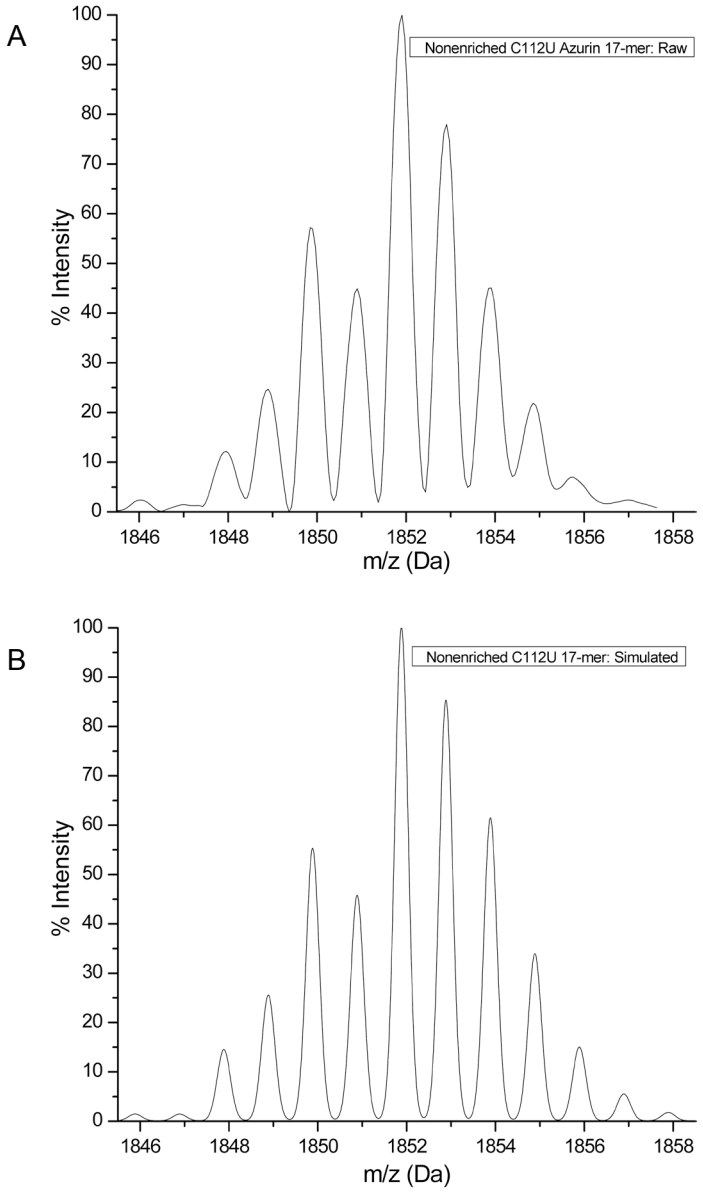


Figure 3.3: Isotopic distribution of non-enriched C112U azurin 17-mer. A) Raw experimental data shows expected distribution for a 17-mer peptide containing a single Se atom. B) Simulated data ($C_{80}H_{133}N_{21}O_{22}SSe$, m/z = 1851.88) using Applied Biosystems Mass Explorer program.

3.5 Optimization and Semi-synthesis of C112U Azurin

Azurin is well suited for EPL because the protein is relatively small at 128 amino acids in length (Figure 3.4). Furthermore, three of the four copper binding ligands are found within the C-terminal 17 amino acids (H_2N-CTFPGHSALMKGTLTLK-COOH) of the protein, making substitutions at these positions conveniently accessible.

AECSVDIQGN DQMQFNTNAI TVDKSCKQFT
VNLSHPGNLP KNVM**GH**NWVL STAADMQGVV
TDGMASGLDK DYLKPDDSRV IAHTKLIGSG
EKDSVTFDVS KLKEGEQYMF F***C****TFPG****H****SAL****M****KGTLTLK*

Figure 3.4: N-terminal to C-terminal amino acid sequence of *Pseudomonas aeruginosa* azurin. **Bold** residues are those involved in copper binding. The final 17 amino acids synthesized by SPPS are *italicized*.

To accomplish incorporation of selenocysteine into azurin using EPL, the procedures established by Berry were followed.[8] The first 111 amino acids of azurin were recombinantly expressed as an Intein-Chitin binding domain (CBD) fusion using NEB vector PtxB1 intein splicing technology. Attempts to isolate the recombinant azurin-thioester were unsuccessful due to its insolubility and facile thioester hydrolysis. Attempts at solubilizing the precipitated protein with guanidinium hydrochloride (GdmCl) were also unsuccessful when attempting to run EPL reactions. Therefore ligations were performed on column with the final 17 amino acid peptide in the presence of an external thiol catalyst.

Initial EPL reactions were completed in TRIS (tris(hydroxymethyl)aminomethane) buffer. However, after exchanging buffers to MOPS (3-(N-morpholino)propanesulfonic acid), large amounts of white precipitate, attributable

to hydrolyzed thioester, as well as full length (FL) product were present. To test both the performance of different buffers during the ligation step and the optimal pH, ligation experiments were performed in various buffers at various pH values. HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), MOPS, TRIS, tricine (N-(tri(hydroxymethyl)methyl)glycine), and potassium phosphate buffers were tested at 50 mM concentrations at pH 7.2 and 8.0 (Figure 3.5). EPL reactions performed in HEPES buffer at pH 8 yielded the most ligated protein and the least amount of precipitate following buffer exchange (Figure 3.5). Interestingly, close inspection of the byproducts of EPL reactions that were performed in TRIS buffer revealed peptide masses consistent with the sequential addition of TRIS by MALDI-MS. Calculations show that the TRIS ligand may be adding into Lys residues and interferes with trans-selenoesterification. Therefore, EPL reactions were completed in HEPES buffer at pH 8.

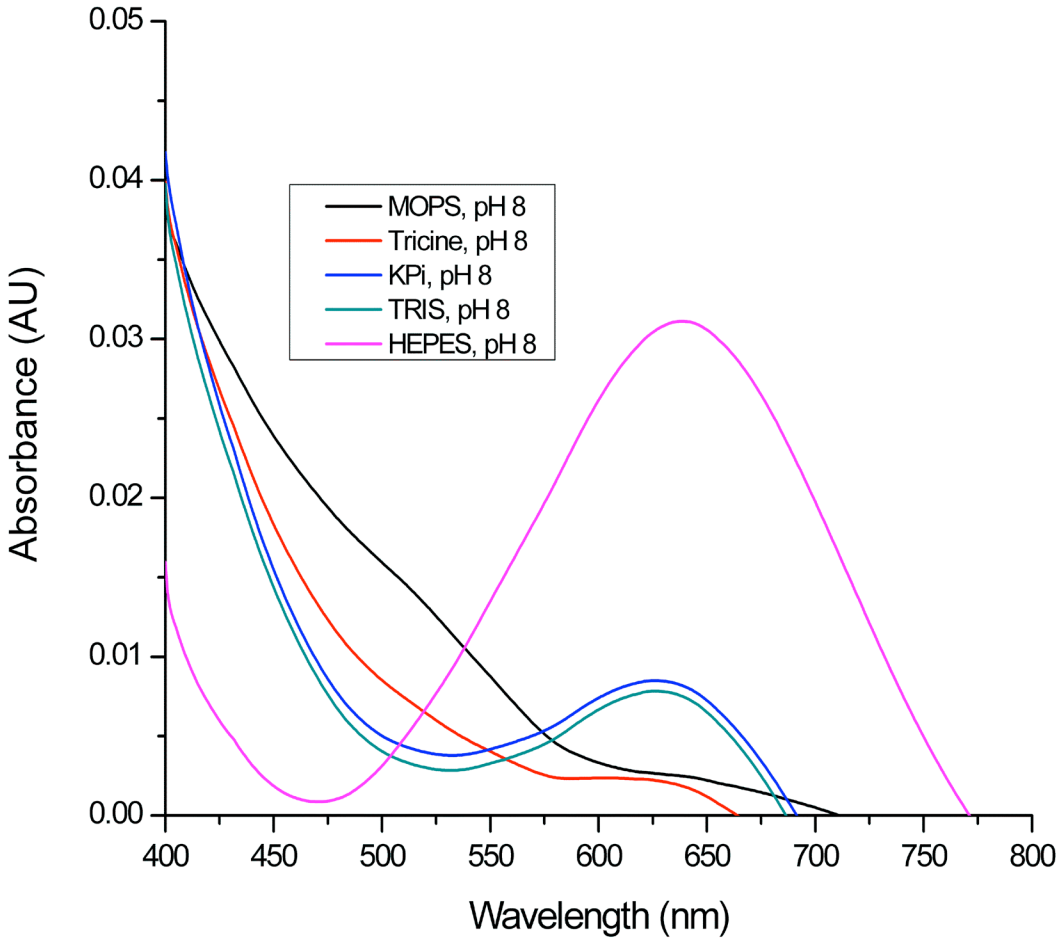


Figure 3.5: EPL performed in differing buffers at pH 8.0. HEPES buffer displays the best RL ratio.

The ratio of the A_{280} and A_{675} peaks, RL [=ε(high energy)/ε(low energy)], provides a means to measure the contribution of the LMCT and serves as a crude method to determine Cu binding efficiency.[30] Even in the optimized buffer system, the RL value of C112U (RL = 15) compared to WT azurin (RL = 2) was low, suggesting the protein inefficiently bound Cu, even though the C112U azurin variant has a lower extinction coefficient than WT azurin.[9] Decreased Cu binding can be attributed to poor protein folding. The folding of azurin has long been studied and debated throughout the

literature,[31-38] but one facet clearly aids in determining how well the protein is folded. A buried tryptophan residue (Trp48) located at the β turn within the azurin structure displays an extraordinary pronounced absorbance at 291 nm when the protein is correctly folded.[33] This is due to the rigidity from the hydrophobic area of the protein that the Trp residue is buried within. When the Cu is correctly bound, and the protein is folded correctly, the A_{280} shows a sharp peak at A_{291}.[33] However, if the protein is incorrectly folded or conformational heterogeneity of the active site exists, this sharp peak is absent.

The apo-C112U Az A_{280} peak showed a broad A_{291} peak, suggestive of poor folding which ultimately contributes to poor Cu binding. To increase proper protein folding apo-C112U azurin was refolded in the presence of GdmCl and DTT (dithiothreitol), a process used in the Lu laboratory to refold WT azurin that typically results in higher RL values for mutants that are poorly folded. When used for C112U azurin, the procedure resulted in complete loss of the A_{280} and no protein with the C112U mutant. The protein degradation is likely a result of oxidative damage to the protein scaffold and deselenization of the Sec112 residue in the presence of DTT. The C112U protein may be more susceptible to both inefficient ligation and improper protein folding as the selenol is much easier to oxidize than the free thiol, supporting why the C112U protein is unable to be properly refolded.

Though unable to solve the refolding issue via a refolding step, another method to reduce folding issues was to perform ligations anaerobically under an Ar atmosphere. The EPL reactions were run in an Ar purged column under an Ar atmosphere for 66 h before elution under Ar. Performing the ligations in an anaerobic environment increased

protein yields (0.32 mg/L growth) as well as the amount of correctly folded protein.

Having now established a new buffer condition (HEPES, pH 8) and that ligations required an anaerobic environment to provide better yields of correctly folded protein, the only remaining variable to further increase ligation efficiency was the thiol catalyst. Kent demonstrated that the leaving group reactivity of the thiol catalyst balanced with its thiophilicity are important for efficient transthioesterification.[39] Thus alkyl-thiols are inferior to aryl-thiols because they are poor leaving groups. Aryl-thiols with pKa values higher than 6.5 selectively participate in transthioesterification. To test whether MESNa (sodium 2-mercaptoethanethiosulfonate) used in ligations discussed thus far, was an inferior thiol catalyst choice in C112U azurin EPL reactions, four alternative thiols, N-methylmercaptoacetamide (NMA), 4-methylbenzenethiol (4-MeBzlthiol), 4-methoxybenzenethiol (4-MeOBzlthiol), and 4-mercaptophenylacetic acid (MPAA) were tested for their ability to increase the yield of the obtained product. NMA, 4-MeBzlthiol, and 4-MeOBzlthiol have prohibitive smells whereas MPAA has no discernable smell. Neither NMA nor 4-MeOBzl thiol resulted in appreciable increases in protein yields, likely due to NMA's poor leaving group efficiency and 4-MeBzlthiol's low solubility in water. No ligation was noted in ligations using the MeBzlthiol likely due to its lower solubility in water than 4-MeOBzlthiol. However, using MPAA resulted in significant yield increases (0.42 mg/L growth) and was used exclusively thereafter.

3.6 CHARACTERIZATION OF SELENOCYSTEINE AZURIN

The optimized procedure using HEPES buffer, pH 8, anaerobic conditions, and MPAA as the thiol catalyst resulted in a C112U azurin mutant that was spectroscopically identical to the previously reported protein but with better yield (~0.4 mg/L cell growth) than the previously reported yield (~0.25 mg/L cell growth) (Figure 3.6).[8,9] With the optimized procedures in place, semi-synthesis of a ^{77}Se enriched C112U variant was undertaken.

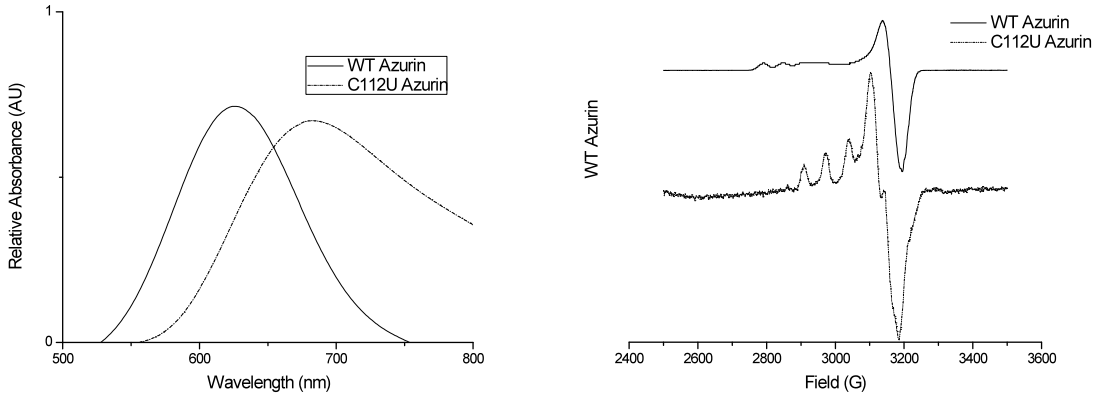


Figure 3.6: Spectroscopic characterization of C112U azurin compared to WT azurin A) UV-Vis spectrum B) EPR spectrum.

3.7 SYNTHESIS OF ^{77}SE-SEC(PMB) AND SPPS OF ^{77}SE-C112U 17MER.

Introducing ^{77}Se enriched Sec into azurin can provide insight into the location of the electron density of the free electron in the cuprice type I site. Laboratories have debated for decades about the distribution of the free electron on the bound Cu(II) in azurin.[5,40-42] Recent DFT calculations using crystal structures of azurin, its mutants, and some relatives, suggested that approximately 27% of the electron density resides on the

Cu and 60% resides on the Cys(S) ligand.[15] The remaining 3% is scattered through the His ligands.[40] Using ENDOR spectroscopy and incorporating a ½ spin labeled ^{77}Se-Sec ligand in the azurin type I site could provide precisely calculated electron contributions for the Cu and the Sec(Se) ligand.

Fmoc-^{77}Se-Sec(PMB)-OAll was previously synthesized through a collaborative effort between the van der Donk laboratory and Pete Silks at Los Alamos National Laboratory. Only the final allyl deprotection step was necessary. Therefore, the optimized allyl deprotection strategy from Scheme 1 was used to afford the enriched Fmoc-^{77}Se-Sec(PMB)-OH derivative in a 90% yield. The optimized SPPS techniques outlined for the non-isotopic enriched sample where followed to make peptide ***3.20*** and the PMB group was removed using DTNP to afford the free selenol peptide ***3.21***. The peptide was analyzed for the correct monoisotopic mass and simulations were run using the Applied Biosystems Data Explorer program (version 4) to simulate the isotopic distribution pattern for the 17-mer peptide containing a ^{77}Se enriched Sec atom ($C_{80}H_{133}N_{21}O_{22}S^{77}Se$, Exact Mass: 1848.89). Figure 3.7 shows that the synthesized peptide has the correct monoisotopic mass and the correct isotopic distribution pattern expected for a peptide containing one ^{77}Se enriched selenium atom.

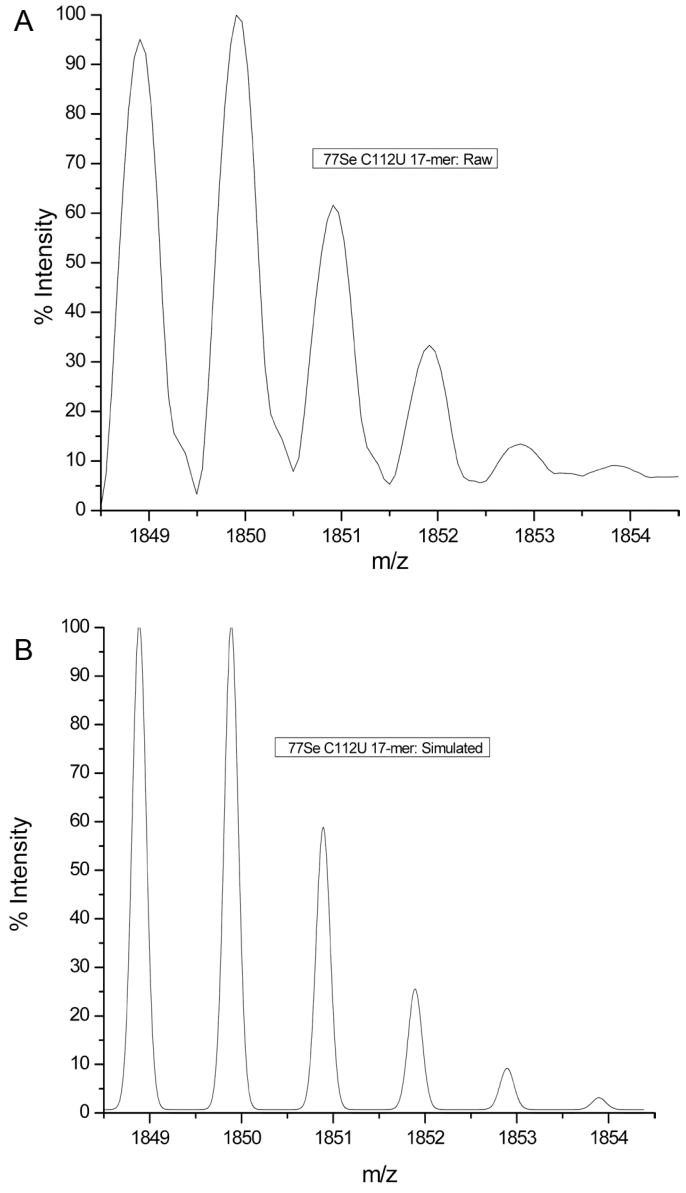


Figure 3.7: Isotopic distribution of ^{77}Se-enriched C112U azurin 17-mer. A) Raw experimental data shows expected distribution for a 17-mer peptide containing a single ^{77}Se atom. B) Simulated data ($C_{80}H_{133}N_{21}O_{22}S^{77}Se$, Exact Mass: 1848.89) using Applied Biosystems Mass Explorer program.

3.8 SEMI-SYNTHESIS OF ^{77}SE-C112U AZURIN

The optimized EPL conditions developed for non-isotopically enriched C112U azurin were used for the incorporation of the enriched ^{77}Se-Sec into azurin. Ligations proceeded for 66 h on the chitin column and were eluted, concentrated, and titrated with Cu(II). As shown in Figure 3.8, the ^{77}Se-C112U variant bound copper in an acceptable RL ration of 3.5. The LMCT band remained red-shifted from WT azurin at 675 nm (Figure 3.8).

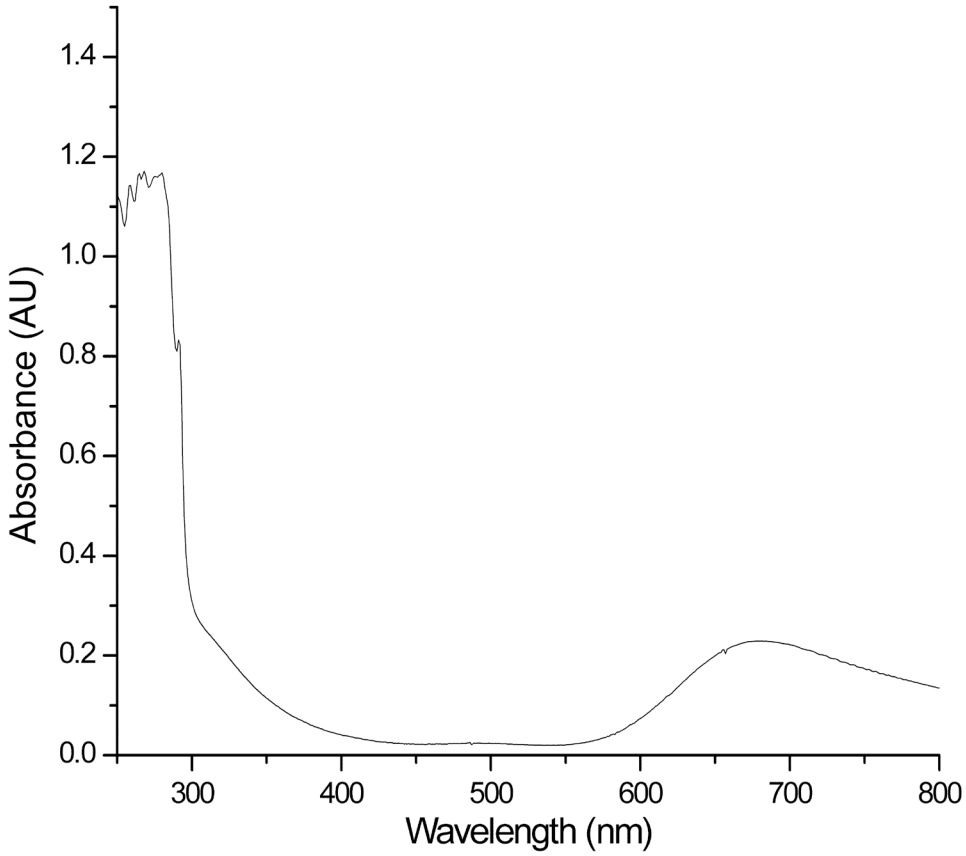


Figure 3.8: UV-Vis spectrum for ^{77}Se C112U azurin.

With the semi-synthesis and purification of the ^{77}Se-C112U azurin protein complete, X-Band EPR was performed before shipping the sample to Dr. David Britt's laboratory at the University of California, Davis for ENDOR spectroscopy (Figure 3.9). The ^{77}Se-C112U azurin spectrum displayed the same *g* values as the non-enriched protein but had more features along $g_{\perp}$, suggesting that the ½ spin of the ^{77}Se atom may be altering the spectrum.[8,9]

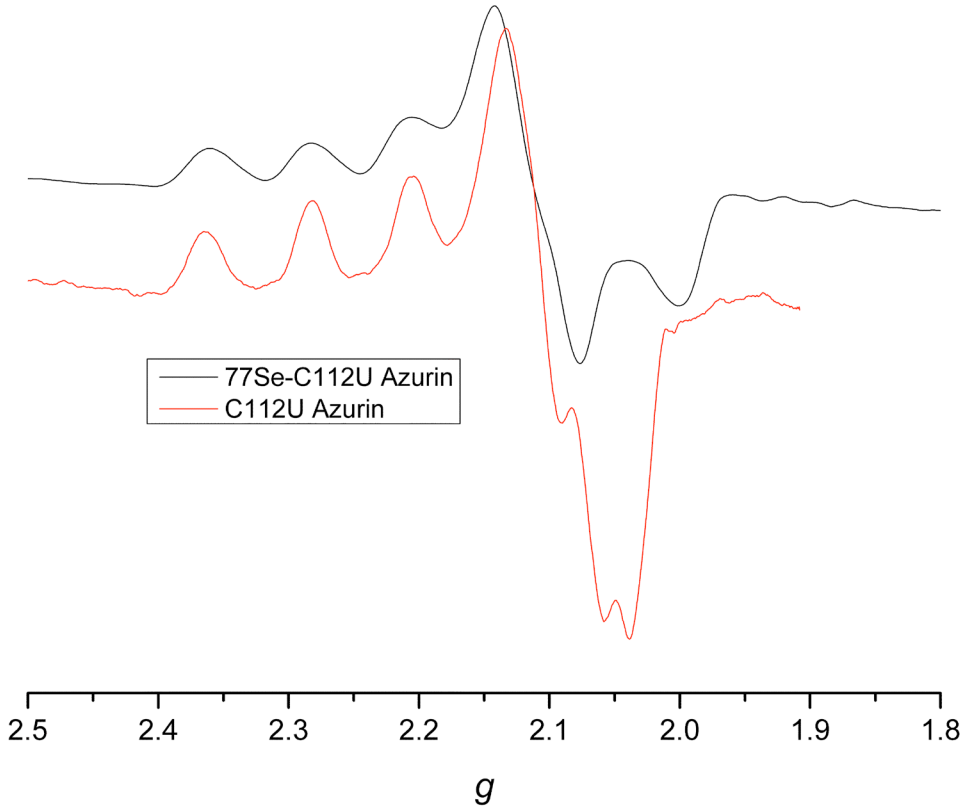


Figure 3.9: X-band EPR spectrum for ^{77}Se C112U azurin.

3.9 ^{77}Se C112U Azurin ENDOR Analysis

The ^{77}Se-C112U protein was shipped to David Britt's laboratory with the intention of obtaining clean Electron-Nuclear Double Resonance (ENDOR) spectroscopy corroborating electron contributions of the Cu and the Sec(Se) ligand. However, no simulation parameters were available for a ^{77}Se atom and had to be created before any simulations could be performed. Once the parameters were created, initial ENDOR spectra showed narrow nitrogen couplings (Figure 3.10). Nitrogens in the WT and nonenriched sample have quadrupolar couplings that broaden the ENDOR signal beyond detection while the nitrogens in the ^{77}Se mutant do not broaden the signal and are detectable.

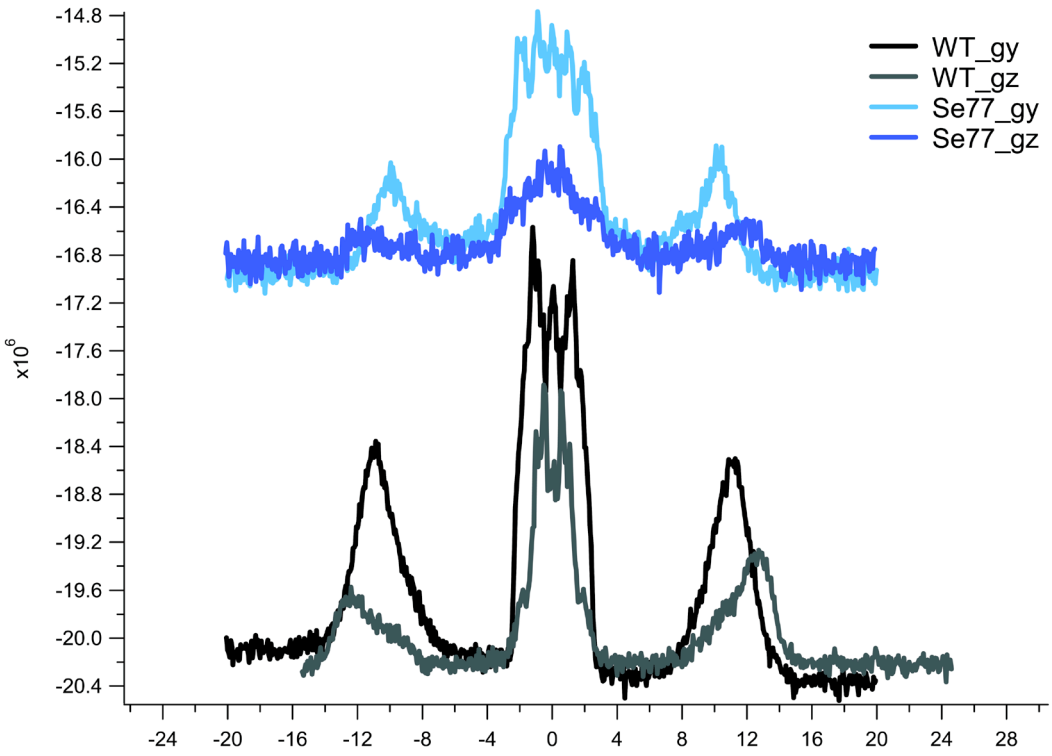


Figure 3.10: Proton Q-band ENDOR. Larmor frequency set to zero.

After struggling to find a region for the ^{77}Se signal, EPR power dependency on the ^{77}Se C112U protein revealed an unusual contribution to the signal on the high field end (Figure 3.11 blue arrow). The feature grows in with increased power, and the signal appears saturated. However, after running ENDOR in the expected region of the ^{77}Se hyperfine tensor, 1-150 MHz, no signal was detectable. Numerous setbacks, including instrumentation failure, cavity contamination, and insufficient concentrations of protein have thus far limited progress in providing evidence for the electron contributions of Cu and Sec(Se) in azurin.

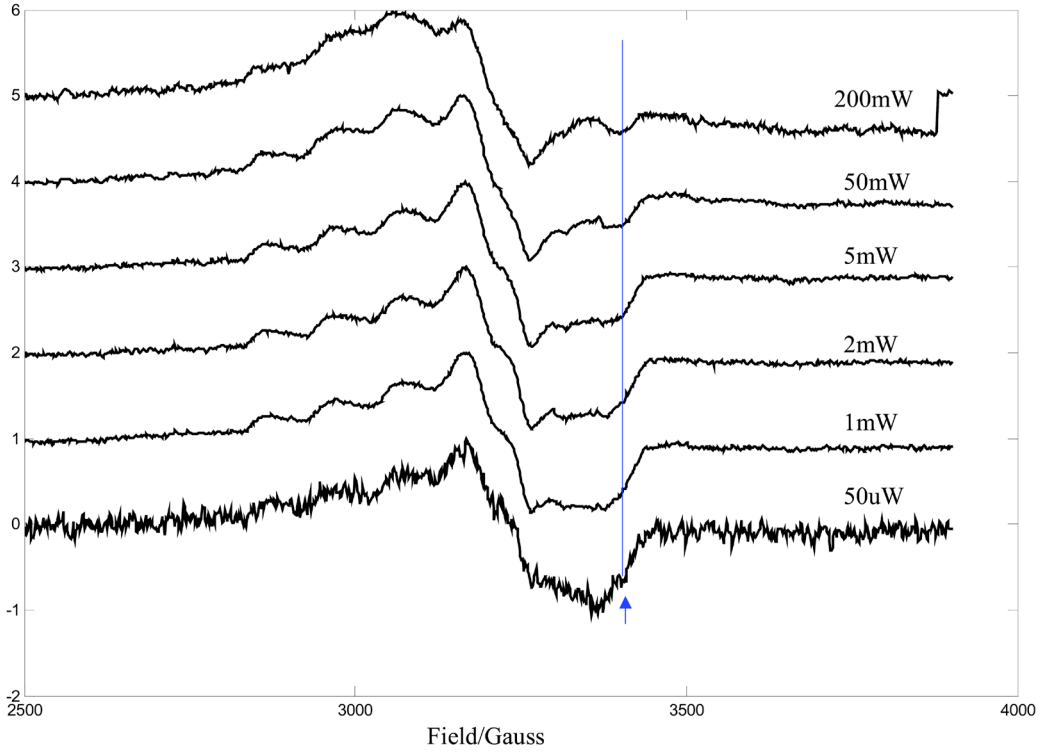


Figure 3.11: ENDOR preparation. EPR power dependency for ^{77}Se C112U Az. The blue arrow indicates the location of the unknown growing feature.

3.10 CONCLUSIONS

Optimizing the incorporation of Sec into azurin using EPL resulted in significant yield increases of holo-protein (0.42 mg/L from 0.2 mg/L).[8,9] The newly optimized synthetic schemes for synthesis of the Sec(PMB) derivative and the SPPS of the 17-mer peptide were used to synthesize a ^{77}Se enriched C112U 17-mer peptide in high yields (60-80 mg). The same optimized EPL reactions for the non-enriched C112U azurin variant were used to semi-synthesize a ^{77}Se-C112U azurin variant that displayed similar spectra to the non-enriched species. However, using the enriched holo-protein has yet to reveal the electron contribution of the Cu and the Sec(Se) ligand due to unforeseen complications in the laboratory of our ENDOR collaborator.

3.11 FUTURE DIRECTIONS

New enriched protein is necessary in a higher concentration to hopefully provide sufficient material for ENDOR. Conditions that identify the ^{77}Se signal in the ENDOR spectra are absolutely necessary before continuing with the project. If parameters cannot be established to identify and simulate ^{77}Se signals to compare with both WT azurin and the non-enriched C112U azurin variant, then a new strategy must be adopted to experimentally determine the electron density of the free electron in the azurin type I site.

3.12 EXPERIMENTAL

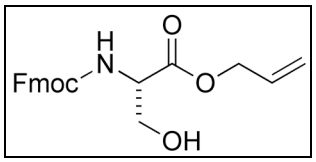
General. All chemicals used were obtained from Aldrich. All amino acids were purchased from Novabiochem, Advanced Chemtech, or Chem-Impex International.

Protecting groups used for Fmoc-SPPS were Cys(Trt), Ser(OtBu), Glu(OtBu), Cys(StBu), Asn(Trt), His(Trt), Lys(Boc), Thr(OtBu). Methylene chloride (DCM) was distilled from calcium hydride. Tetrahydrofuran (THF) was distilled from sodium and benzophenone. All aqueous buffers were degassed under Ar or N_2 as specified immediately prior to use.

All NMR spectra were obtained with either a Varian U400 or U500 NMR Spectrometer in $CDCl_3$. UVVis spectra were taken on an HP Diode Array Spectrometer or a Cary 5000 spectrometer. X-band EPR spectra were collected on a Varian E-122 spectrometer at the Illinois EPR Research Center (IERC). The samples were run at ~30 K using liquid He and an Air Products Helitran cryostat without glassing agents. The magnetic field was calibrated using a Metrolab PT 2025 NMR Gaussmeter, and the microwave frequency was measured with an EIP model 578-frequency counter equipped with a high-frequency option.

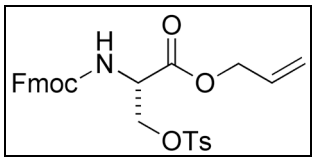
Mass spectral data was collected by the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois by either Electrospray Ionization (ESI MS), or matrix assisted laser desorption ionization (MALDI) techniques as indicated. Mass spectral data are reported as – mass, peak (percent).

Preparative RP-HPLC was performed with a Waters Delta 600 system. Solution A was 0.1 % TFA in H_2O, and solution B was 80 % MeCN/20 % H_2O with 0.1 % TFA. Unless otherwise stated, a linear gradient of 20% to 70% B over 23 min was used for all runs. Solid phase peptide synthesis (SPPS) was performed on a Rainin PS3 automated peptide synthesizer or manually using medium Büchner fritted funnels.[43]



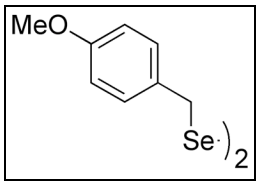
Fmoc-L-Serine-OAll (*3.1*), Lab Book I, page 6.

Fmoc-L-Serine (10.0 g, 30.5 mmol) and $NaHCO_3$ (2.56 g, 30.5 mmol) were dissolved in ddH_2O (46.9 mL) under N_2. Allyl bromide (2.91 mL, 33.49 mmol) and aliquat-336 (12.33 g, 30.5 mmol) were dissolved in CH_2Cl_2 (46.9 mL) and added to the amino acid mixture mixture. The reaction was stirred vigorously at 25 °C under N_2 for 24 h. Allyl bromide (2.91 mL, 33.49 mmol) was added after 24 h of stirring and the reaction was stirred at 25 °C under N_2 for another 48 h. The reaction was extracted with CH_2Cl_2, dried over $MgSO_4$, filtered, and the solvent evaporated. The mixture was purified by silica gel flash chromatography (2:1 hexane/EtOAc) to obtain product as a white solid (R_f = 0.12, 2:1 hexane/EtOAc). (Yield: 10.7 g, 95%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 2.33 (s, 1 H), 3.96 (d, 1 H, *J* = 10.3 Hz), 4.03 (d, 1 H, *J* = 10.1 Hz), 4.23 (t, 1 H, *J* = 6.9 Hz), 4.43 (m, 2 H), 4.48 (m, 1 H), 4.69 (d, 2 H, *J* = 5.3 Hz), 5.26 (dd, 1 H, *J* = 0.8, 10.5 Hz), 5.37 (dd, 1 H, *J* = 0.8, 17.2 Hz), 5.69 (d, 1 H, *J* = 7.4 Hz), 5.91 (ddt, 1 H, *J* = 16.6, 11.0, 5.5 Hz), 7.32 (t, 2 H, *J* = 7.4 Hz), 7.41 (t, 2 H, *J* = 7.4 Hz), 7.60 (m, 2 H), 7.76 (d, 2 H, *J* = 7.6 Hz).



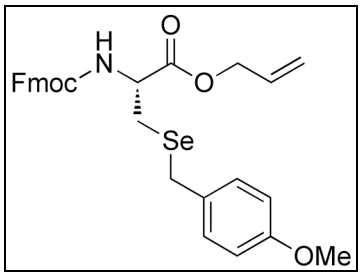
Fmoc-(OTs)-L-Serine-OAll (*3.2*), Lab Book I, page 88.

TsCl (5.19 g, 27.21 mmol) was dissolved in pyridine (11 mL) and stirred at 25 °C for 5-10 min. The reaction mixture was added to solid Fmoc-L-Serine-OAll (2.0 g, 5.44 mmol) under Ar and stirred under Ar for 14 h at 0 °C. The reaction was quenched with EtOAc (50 mL) and the organic layer was washed with ddH_2O, 10% $KHSO_4$, saturated $NaHCO_3$, and brine. The organic layer was dried over $MgSO_4$, filtered, and the solvent evaporated. The reaction was purified quickly by silica gel flash chromatography (3:1 hexane/EtOAc) to obtain a clear oil. The oil was azeotroped with CH_2Cl_2 (3 x 50 mL) and with Et_2O (3 x 50 mL) to obtain product as a white solid. (R_f = 0.19, 3:1 hexane/EtOAc). (Yield: 4.2 g, 79%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 2.36 (s, 3 H), 4.19 (t, 1 H, *J* = 7.5 Hz), 4.27 (dd, 2 H, *J* = 10.2, 7.2 Hz), 4.36 (m, 2 H), 4.45 (dd, 2 H, *J* = 10.2, 3.0 Hz), 4.61 (m, 2 H), 4.67 (m, 1 H), 5.27 (dd, 2 H, *J* = 10.4, 0.9 Hz), 5.33 (dd, 1 H, *J* = 17.2, 1.0), 5.65 (d, 1 H, *J* = 7.7 Hz), 5.84 (m, 1 H), 7.26 (t, 2 H, *J* = 8.7 Hz), 7.33 (m, 2 H), 7.42 (t, 2 H, *J* = 7.1, 2.2 Hz), 7.60 (t, 2 H, *J* = 8.4 Hz), 7.77 (m, 4 H).



Di-p-methoxybenzyl diselenide ($(PMBSe)_2$) ***(3.3)***. Lab Book I, page 80.

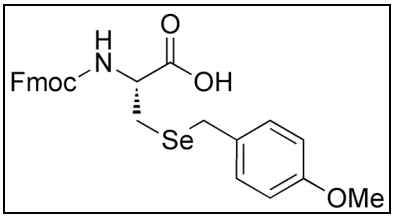
Selenium powder (2.0 g, 25.3 mmol) was suspended in dry THF (12 mL) and purged with Ar. The mixture was brought to a reflux prior to adding 1 M $LiB(Et_3)H$ (25.1 mL, 25.1 mmol) dropwise. The mixture was refluxed for 45 min under Ar. In a separate round bottom flask, *p*-methoxybenzylchloride (PMBCl) (3.62 mL, 26.6 mmol) was dissolved in dry THF (12 mL) and cooled to -78 °C. The refluxing mixture was cooled to 25 °C and transferred to the cold flask containing PMBCl via cannulation. The mixture was stirred at -78 °C for 45 min and then slowly warmed to 25 °C over 1 h. The reaction mixture was quickly run through a silica gel plug using 5:1 hexane/EtOAc and purified by silica gel flash chromatography (9:1 hexane/EtOAc) to obtain yellow solid which was recrystallized in EtOH to obtain final product as light yellow solid. (R_f = 0.5, 5:1 hexane/EtOAc). (Yield: 3.99 g, 81%). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 3.80 (s, 6 H), 3.84 (s, 4 H), 6.83 (d, 4 H, J = 8.4), 7.17 (d, 4 H, J = 8.4).



Fmoc-(Se)-*p*-methoxybenzylselenocysteine-OAll (*3.4*), Lab book II, page 36.

A three-necked round bottom flask was charged with compound ***3.3*** (2.56 g, 6.39 mmol), purged with Ar and dissolved in dry THF (15 mL). The mixture was brought to a reflux prior to adding 50% H_3PO_2 (23 mL) dropwise. The reaction was then refluxed for an additional 1 h. The refluxing mixture was cooled to ambient temperature, quickly extracted with Et_2O, and washed with brine. The combined organic layers were dried with $MgSO_4$, filtered, and the solvent removed to obtain a yellow oil. The resultant selenol was dissolved in degassed DMF (15 mL), purged with Ar, and cooled to 0 ºC. Degassed 1 M NaOH in ddH_2O (4.5 mL, 4.5 mmol) was added to the flask followed immediately by compound ***3.2*** (2.1 g, 4.03 mmol) dissolved in degassed acetone (15 mL). The mixture was stirred at 0 °C for 2-4 h. The reaction was taken up in EtOAc (50 mL) and washed with saturated NH_4Cl and brine (3 x 75 mL each). The organic layer was dried over $MgSO_4$ and the solvent evaporated. The reaction was purified by silica gel flash chromatography (6:1 hexane/EtOAc) to obtain final product as white solid. (R_f = 0.29, 6:1 hexane/EtOAc). (Yield: 1.85 g, 84%). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 2.94 (1H, B of A_{BX}, J_{AB} = 13.0 Hz, J_{BX} = 5.5 Hz), 3.77 (s, 5 H), 4.24 (t, 1 H, J = 6.8 Hz), 4.42 (d, 2 H, J = 6.8 Hz), 4.66 (d, 2 H, J = 5.6 Hz), 4.7 (1H, X of A_{BX}, J_{BX} = 5.5 Hz, J_{AX} = 4.4 Hz), 5.26 (dd, 1 H, J = 10.4, 1.1 Hz), 5.36 (dd, 1 H, J = 17.2, 1.0 Hz), 5.55 (d, 1 H, J = 7.9

Hz), 5.90 (m, 1 H), 6.83 (d, 1 H, *J* = 8.4 Hz), 7.20 (d, 2 H, *J* = 8.4 Hz), 7.31 (m, 2 H), 7.40 (m, 2 H), 7.61 (t, 2 H, *J* = 6.2 Hz), 7.77 (dd, 2 H, *J* = 7.8, 2.1 Hz).



Fmoc-Se-p-methoxybenzylselenocysteine-OH (*3.5*), Lab book II, page 41.

$Pd(PPh_3)_4$ (31.4 mg, 0.027 mmol) and PPh_3 (8 mg, 0.027 mmol) were combined in a dried round bottom flask and purged with Ar. The flask was covered with aluminum foil and CH_2Cl_2 (5 mL) was added. Na-2-ethylhexanoate (181 mg, 1.09 mmol) dissolved in CH_2Cl_2 (5 mL) was added to the Pd catalyst and the reaction was stirred vigorously at 25 °C for 15 min under Ar. The reaction was taken up in EtOAc (50 mL) and washed with 2 M HCl and 4 M HCl (2 x 50 mL each). The organic layer was dried over $MgSO_4$ and the solvent evaporated to obtain a yellow oil. The mixture was purified by silica gel flash chromatography (35:1 CH_2Cl_2/MeOH with 0.1% AcOH (v/v)) to obtain final product as white solid. (R_f = 0.5, 9:1 CH_2Cl_2/MeOH). (Yield: 430 mg, 93%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 2.96 (br s, 2 H), 3.75 (s, 3 H), 3.76 (br s, 2 H), 4.23 (t, 1 H, *J* = 6.8 Hz), 4.42 (d, 2 H, *J* = 6.8 Hz), 4.67 (m, 1 H), 5.52 (d, 1 H *J* = 7.2 Hz), 6.80 (d, 2 H, *J* = 8.3 Hz), 7.18 (d, 2 H *J* = 8.1 Hz), 7.30 (t, 2 H, *J* = 7.4 Hz), 7.39 (m, 2 H), 7.60 (t, 2 H, *J* = 6.7 Hz), 7.76 (d, 2 H *J* = 7.6 Hz).

H_2N-$U_{(PMB)}$TFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.10*, Lab book IV, page 178.

The peptide was synthesized on a 0.1 mmol scale. Fmoc Wang resin preloaded with Fmoc-Lys(Boc) was pre-swelled in DMF (6 x 10 min, 6 mL). Synthesis of the peptide from Lys128 through Thr113 was completed using a four-fold excess of amino acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (4 equiv) for 2 min prior to coupling to the resin. All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-Sec(PMB)-OH (***3.5***) (4 equiv.) was coupled manually in the presence of 1-hydroxybenzotriazole (HOBt) (8 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (8 equiv.) in DMF at room temperature (RT) for 2-5 h. The Fmoc was removed with 20% piperidine and the resin washed with DCM before cleavage. Amino acids were double coupled when necessary and couplings were monitored by qualitative ninhydrin test. (Yield: 90 mg). MALDI MS: Expected. 1972.2 Observed. 1972.6.

H_2N-UTFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.11* Using I_2 in AcOH, Lab book 4, page 106.

Peptide ***3.10*** (40 mg, 0.02 mmol) was placed into a round bottom flask and purged with N_2. I_2 (6 mg, 0.02 mmol) dissolved in 8:1:1 AcOH/MeCN/H_2O (10 mL) was added to the peptide and the reaction was stirred at 25 °C for 5 min. The reaction was

quenched with 10% $Na_2S_2O_3$ (10 mL) and lyophilized overnight before purifying by preparative RP-HPLC to yield the final product as white solid after lyophilization. (Yield: 15 mg). MALDI MS. Expected: 1852.06 (SeH), Observed: 1774.1 (-Se), 1852.06 (SeH), 1972.8 (SePMB); 3:1:2.

H_2N-UTFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.11* using TMSOTf, Lab book 4, page 107.

Peptide ***3.10*** (12 mg, 0.006 mmol) was placed into a round bottom flask and purged with N_2. Then 2.1 mL of 50:120:690:194 *m*-cresol/thioanisole/TFA/TMSOTf was added to the peptide and the reaction was stirred rapidly under N_2 for 30 min. The solvent was removed via rotary evaporation and the crude product was lyophilized. The crude material was purified by preparative RP-HPLC to obtain a white solid that was later identified as deselenated peptide. No product was yielded. MALDI MS. Expected: 1852.06, Observed: 1774.1 (-Se).

H_2N-UTFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.11* Using NpsCl, Lab book 4, page 106.

Peptide ***3.10*** (49 mg, 0.025 mmol) and NpsCl (4.7 mg, 0.25 mmol) were combined in a round bottom flask and dissolved in 1:1 MeCN/H2O. The reaction was initiated by adding AcOH (3 mL) and 2 M HCl (75 μL) and stirred at 25 °C for 30 min. The solvent was removed by rotary evaporation and the peptide was triturated with cold

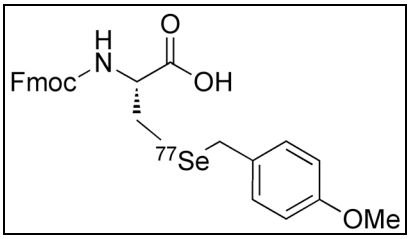
Et_2O. The peptide was dissolved in 1:1 MeCN/H_2O (10 mL) and lyophilized prior to purification by preparative RP-HPLC to obtain a yellow solid. (Yield: 35 mg). MALDI MS. Expected: 1852.06 (SeH), 2006.6 (SeNps), Observed: 1852.06 (SeH), 2006.6 (SeNps); 1:2.

H_2N-UTFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.11* Using DTNP, Lab book IV, page 178.

Resin bound, Fmoc-deprotected peptide ***3.10*** (0.1 mmol) was washed with DCM and dried under reduced pressure for 10 min. The resin was transferred to a scintillation vial and 2,2'-dithiobis(5-nitropyridine) (DTNP) (1.3 mol equiv., 41 mg) was added. 2% (v/v) triisopropylsilane (100 μL) and 2% H_2O (100 μL) were added followed by the addition of 2.5% (v/v) thioanisole (125 μL). Cleavage was initiated by the addition of 5 mL of neat TFA and the reaction was stirred at 25 °C for 2 h. The reaction was filtered through a coarse grain Büchner funnel and the resin was washed with neat TFA (3 x 5 mL). The solvent was removed via rotary evaporation to obtain a yellow oil. The peptide was precipitated by the addition of cold Et_2O and pelleted by centrifugation at 2000 x g for 5 min. The pellet was washed iteratively with Et_2O (3 x 30 mL) with centrifugation steps between each washing. The peptide was dissolved in a minimal amount of 1:1 MeCN/H_2O and lyophilized overnight. The crude lyophilized product was purified by RP-HPLC C18 column to obtain final product as yellow solid. (Yield: 81 mg). MALDI

MS. Expected: 1852.06 (SeH), 2006.21 (Se(pys)), 3704.1 (Dimer), Observed: 1853.18 (M+1H), 3703.69 (Dimer).



Fmoc-^{77}Se-p-methoxybenzylselenocysteine-OH (*3.7*), Lab book V, page 8.

A round bottom flask was charged with $Pd(PPh_3)_4$ (21.07 mg, 0.05 mmol) and PPh_3 (4.72 mg, 0.05 mmol) and purged with Ar. A second round bottom flask was charged with Fmoc-(^{77}Se)-*p*-methoxybenzylselenocysteine-OAll (***3.6***) (200 mg, 0.365 mmol) and 1.3 equiv. sodium 2-ethylhexanoate (61.22 mg, 0.368 mmol) and purged with Ar. The contents of each flask were dissolved in dry THF (5 mL). The contents of the amino acid flask were transferred to the Pd catalyst via cannulation under Ar pressure. The reaction was covered with foil and stirred vigorously at 25 °C for 15-20 min. The reaction was quenched with EtOAc (50 mL) and the organic solution was washed with 2 M, 4 M, and 6 M HCl (2 x 40 mL each). The organic layer was dried over $MgSO_4$ and the filtrate stirred gently with PS-PPh_3 resin at 25 °C for 10-20 min. The organics were filtered from the PS-PPh_3 resin using a coarse fritted Büchner funnel and the resin was washed with EtOAc (3 x 20 mL). The solvent was removed by rotary evaporation to obtain a yellow oil. The oil was purified by silica gel flash chromatography (35:1 CH_2Cl_2/MeOH with 0.1% AcOH (v/v)) to obtain final product as white solid. (R_f = 0.5,

9:1 CH_2Cl_2/MeOH). (Yield: 164.2 mg, 89%). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 2.94 (d, 2 H), 3.75 (s, 2 H), 3.77 (br s, 2 H), 4.23 (t, 1 H, *J* = 6.8 Hz), 4.41 (d, 2 H, *J* = 6.8 Hz), 4.67 (m, 1 H), 5.51 (d, 1 H, *J* = 7.2 Hz), 6.81 (d, 2 H, *J* = 8.3 Hz), 7.17 (d, 2 H, *J* = 8.1 Hz), 7.30 (t, 2 H, *J* = 7.4 Hz), 7.39 (m, 2 H), 7.60 (t, 2 H, *J* = 6.7 Hz), 7.76 (d, 2 H, *J* = 7.6 Hz).

H_2N-$^{77}U_{(PMB)}$TFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.20*, Lab book V, page 9.

The peptide was synthesized on a 0.1 mmol scale. Fmoc Wang resin preloaded with Fmoc-Lys(Boc) was pre-swelled in DMF (6 x 10 min, 6 mL). Synthesis of the peptide from Lys128 through Thr113 was completed using a five-fold excess of amino acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (5 equiv) for 2 min prior to coupling to the resin. All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-^{77}Se-Sec(PMB)-OH (***3.7***) (3 equiv.) was coupled manually in the presence of hydroxybenzotriazole (HOBt) (6 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (6 equiv.) in DMF at room temperature (RT) for 2-5 h under N_2 sparging. The Fmoc group was removed with 20% piperidine and the resin washed with DCM before cleavage. Amino acids were double coupled when necessary and couplings were monitored by qualitative ninhydrin test. (Yield: 50 mg).

H_2N-^{77}UTFPGHSALMKGTLTLK-COOH

Synthesis of Peptide *3.21*, Lab book V, page 9.

Resin bound, fully protected peptide was washed with DCM and dried under reduced pressure for 10 min. The resin was transferred to a scintillation vial and 2,2'-dithiobis(5-nitropyridine) (DTNP) (1.3 mol equiv., 41 mg) was added to the resin. 2% (v/v) triisopropylsilane (100 μL) and 2% H_2O (100 μL) were added to the scintillation vial followed by the addition of 2.5% (v/v) thioanisole (125 μL). Cleavage was initiated by the addition of 5 mL of neat TFA and the reaction was stirred at 25 °C for 2 h. The cleavage mixture was filtered through a coarse grain Büchner funnel and the resin was washed with neat TFA (3 x 5 mL). The solvent was removed by rotary evaporation to obtain a yellow oil. The peptide was precipitated by the addition of cold Et_2O. The peptide was pelleted by centrifugation at 2000 x g for 5 min and the pellet was washed iteratively with Et_2O (3 x 30 mL) with centrifugation steps between each washing. The peptide was then dissolved in a minimal amount of 1:1 MeCN/H_2O and lyophilized overnight. The crude lyophilized product was purified by RP-HPLC C18 column to obtain final product as yellow solid. (Yield: 50 mg). MALDI MS; Calculated. 1850.06 (SeH), 2004.21 (Se(pys)), 3700.1 (Dimer) Observed. 1850.29 (M), 3698.9 (Dimer).

General procedure for expressed protein ligation for nonenriched C112U Azurin, Lab Book VI, page 174.

Cultures of *E. coli* BL21 (DE3) cells containing a plasmid to express the azurin(1-111)-Intein-CBD fusion protein were grown in LB media for 8 h at 37 ºC and were used to inoculate eight 2 L flasks of LB media containing 100 mg/mL ampicillin. The cells were grown at 37 ºC for 16 h with shaking at 210 rpm. Protein expression was induced at ~16 h with 0.3 mM IPTG and induction was continued for 4 h at 37 ºC. The cells were then harvested at 9800 x g, stored at -20 ºC and used when needed.

Frozen azurin-thioester cell stock was re-suspended in a lysis buffer containing 20 mM HEPES, pH 7.2, 250 mM NaCl, 1 mM EDTA, 1 mM PMSF, and 0.1% Triton-X-100. The crude lysate was by lysed sonication (9 min work time: 6 s on, 10 s rest). After the second pass, the lysate was treated with 2 M urea for 10 min with gentle stirring at 4 °C. The lysate was centrifuged at 13250 x g for 30 min. at 4 °C. The fusion protein was then bound by batch absorption to 70 mL of chitin resin that had been pre-equilibrated with 20 mM HEPES, pH 8.0, 250 mM NaCl, and 1 mM EDTA (buffer 1) for 1-2 h at 4 ºC. The chitin resin was then poured into a column and the column headspace was purged with Ar. The column was then washed with 3 column volumes of degassed buffer 1 containing 2 M urea under Ar pressure followed by 5-7 column volumes degassed buffer 1.

Ligation was initiated by the addition of peptide (0.52 mM, 50 mg) and tris-(2-carboxyethyl)phosphine (TCEP) (1.2 mM, 17 mg) in 50 mL of buffer 1 containing 50 mM sodium 2-mercaptoethansulfonate (MESNa) under Ar directly to the column. The

chitin resin was then re-suspended in the column and the entire column was agitated gently at 4 ºC for 66 h.

After ligation, the column was eluted under Ar pressure and washed with 1 column volume of buffer 1. The eluent was centrifuged at 13250 x g for 30 min and the supernatant was concentrated 5x using 10,000 MWCO Centricon concentration spin tubes at 2000 x g and the buffer exchanged to 50 mM MOPS, pH 6.6 via G25 sephadex PD-10 columns. The desalted sample was then titrated with incremental additions of freshly made $CuSO_4$ solution (20 mM) in ddH_2O that had been pretreated with chelex. The titration was performed at 4 ºC until maximal absorbance at 675 nm was reached. Samples were then purified via FPLC using anion exchange chromatography. The samples were purified using a Pharmacia Q HiTrap HP column (5 mL) which had been washed with 10 column volumes of buffer B (A: 50 mM MOPS, pH 6.6; B: 1 M NaCl in H_2O) and then equilibrated with 10 column volumes of buffer A.

General procedure for expressed protein ligation for ^{77}Se enriched C112U Azurin, Lab Book VII, page 94.

The ligations were carried out identically to the non-isotopically enriched samples with the following changes. The ^{77}Se-Sec 17-mer peptide (1.3 mM, 71 mg) and tris-(2-carboxyethyl)phosphine TCEP (2.6 mM, 23 mg) were dissolved in 30 mL Buffer 1 containing 50 mM 40-mercaptophenylacetic acid (MPAA) under Ar and transferred to the column under Ar. The column was then shaken gently at 4 °C for 66 h. After ligation, the column was eluted under Ar pressure and washed with 1 column volume of Buffer 1.

The eluent was centrifuged at 13250 x g for 30 min and the supernatants were combined and concentrated using 10,000 MWCO centricon concentration spin tubes at 2000 x g. The buffer was exchanged to 50 mM MOPS Buffer, pH 6.6 via PD10 (G25 sephadex) columns. The desalted sample was then concentrated to 2 mL and purified by a S-100HR SEC column (20 mL bed volume). The fractions containing protein were pooled, concentrated and titrated with incremental additions of freshly made $CuSO_4$ solution (20 mM) in ddH_2O that had been pretreated with chelex. The titration was performed at 4 ºC until maximal absorbance at 675 nm was reached. Samples were then purified via a second S-100HR SEC column.

3.13 REFERENCES

(1) Lu, Y. "Metalloprotein and Metallo-DNA/RNAzyme Design: Current Approaches, Success Measures, and Future Challenges." *Chem. Rev.,* **2006**, *45*, 9930-9940.

(2) Lu, Y.; Berry, S. M.; Pfister, T. D. "Engineering Novel Metalloproteins: Design of Metal-Binding Sites into Native Protein Scaffolds." *Inorg. Chem.,* **2001**, *101*, 3047-3080.

(3) Lu, Y. "Design and engineering of metalloproteins containing unnatural amino acids or non-native metal-containing cofactors." *Curr. Opin. Chem. Biol.,* **2005**, *9*, 118-126.

(4) Lu, Y. "Biosynthetic Inorganic Chemistry." *Angew. Chem., Int. Ed,* **2006**, *45*, 5588-5601.

(5) Malmström, B. G. "Rack-Induced Bonding In Blue-Copper Proteins." **1994**, *223*, 711-718.

(6) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Roos, B. O. "Theoretical Study of the Electronic Spectrum of Plastocyanin." *J. Am. Chem. Soc.,* **1997**, *119*, 218-226.

(7) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. "Electronic Structures of Active Sites in Copper Proteins: Contributions to Reactivity." *Chem. Rev.,* **1992**, *92*, 521-542.

(8) Berry, S. M.; Gieselman, M. D.; Nilges, M. J.; van der Donk, W. A.; Lu, Y. "An Engineered Azurin Variant Containing a Selenocysteine Copper Ligand." *J. Am. Chem. Soc.,* **2002**, *124*, 2084-2085.

(9) Ralle, M.; Berry, S. M.; Nilges, M. J.; Gieselman, M. D.; van der Donk, W. A.; Lu, Y.; Blackburn, N. J. "The Selenocysteine-Substituted Blue Copper Center: Spectroscopic Investigations of Cys112SeCys *Pseudomonas aeruginosa* Azurin." *J. Am. Chem. Soc.,* **2004**, *126*, 7244-7256.

(10) Sarangi, R.; Gorelsky, S. I.; Basumallick, L.; Hwang, H. J.; Pratt, R. C.; Stack, T. D. P.; Lu, Y.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. "Spectroscopic and Density Functional Theory Studies of the Blue-Copper Site in M121SeM and C112SeC Azurin: Cu-Se Versus Cu-S Bonding." *J. Am. Chem. Soc.,* **2008**, *130*, 3866-3877.

(11) Moroder, L. "Isosteric Replacement of Sulfur with Other Chalcogens in Peptides and Proteins." *J. Pept. Sci.,* **2005**, *11*, 187-214.

(12) Gieselman, M. D.; Zhu, Y.; Zhou, H.; Galonic, D.; van der Donk, W. A. "Selenocysteine Derivatives for Chemoselective Ligations." *Chembiochem,* **2002**, *3*, 709-716.

(13) Wessjohann, L. A.; Schneider, A.; Abbas, M.; Brandt, W. "Selenium in chemistry and biochemistry in comparison to sulfur." *Biol. Chem.,* **2007**, *388*, 997-1006.

(14) Johansson, L.; Gafvelin, G.; Arner, E. S. J. "Selenocysteine in proteins--properties and biotechnological use." *Biochim. Biophys. Acta.,* **2005**, *1726*, 1-13.

(15) Fittipaldi, M.; Warmerdam, G. C. M.; de Waal, E. C.; Canters, G. W.; Cavazzini, D.; Rossi, G. L.; Huber, M.; Groenen, E. J. J. "Spin-Density Distribution in the Copper Site of Azurin." *ChemPhysChem,* **2006**, *7*, 1286-1293.

(16) Theodoropoulos, D.; Schwartz, I. L.; Walter, R. "Synthesis of Selenium-Containing Peptides." *Biochem.,* **1967**, *6*, 3927-3932.

(17) Theodoropoulos, D.; Schwartz, I. L.; Walter, R. "New synthesis of L-selenocysteine derivatives and peptides." *Tetrahedron Lett.,* **1967**, *8*, 2411-2414.

(18) Walter, R.; Chan, W.-Y. "Syntheses and pharmacological properties of selenium isologs of oxytocin and deaminooxytocin." *J. Am. Chem. Soc.,* **1967**, *89*, 3892-3898.

(19) Gieselman, M. D.; Xie, L.; van der Donk, W. A. "Synthesis of a Selenocysteine-Containing Peptide by Native Chemical Ligation." *Org. Lett.,* **2001**, *3*, 1331-1334.

(20) Jeffrey, P. D.; McCombie, S. W. "Homogeneous, palladium(0)-catalyzed exchange deprotection of allylic esters, carbonates and carbamates." *J. Org. Chem.,* **1982**, *47*, 587-590.

(21) Hondal, R. J.; Raines, R. T. "Semisynthesis of proteins containing selenocysteine." *Methods Enzymol.,* **2002**, *347*, 70-83.

(22) Eckenroth, B.; Harris, K.; Turanov, A. A.; Gladyshev, V. N.; Raines, R. T.; Hondal, R. J. "Semisynthesis and Characterization of Mammalian Thioredoxin Reductase." *Biochem.,* **2006**, *45*, 5158-5170.

(23) Eckenroth, B. E.; Lacey, B. M.; Lothrop, A. P.; Harris, K. M.; Hondal, R. J. "Investigation of the C-Terminal Redox Center of High-Mr Thioredoxin Reductase by Protein Engineering and Semisynthesis." *Biochem.,* **2007**, *46*, 9472-9483.

(24) Hunter, M. J.; Komives, E. A. "Deprotection of S-Acetamidomethyl Cysteine-Containing Peptides by Silver Trifluoromethanesulfonate Avoids the Oxidization of Methionines." *Anal. Biochem.,* **1995**, *228*, 173-177.

(25) Otaka, A.; Koide, T.; Shide, A.; Fujii, N. "Application of dimethylsulphoxide(DMSO) / trifluoroacetic acid(TFA) oxidation to the synthesis of cystine-containing peptide." *Tetrahedron Lett.,* **1991**, *32*, 1223-1226.

(26) Yajima, H.; Fujii, N.; Funakoshi, S.; Watanabe, T.; Murayama, E.; Otaka, A. "New strategy for the chemical synthesis of proteins." *Tetrahedron,* **1988**, *44*, 805-819.

(27) Harris, K. M.; Flemer, S.; Hondal, R. J. "Studies on deprotection of cysteine and selenocysteine side-chain protecting groups." *J. Pept. Sci.,* **2007**, *13*, 81-93.

(28) Hondal, R. J. "Incorporation of Selenocysteine into Proteins Using Peptide Ligation." *Protein Pept. Let.,* **2005**, *12*, 757-764.

(29) Kudryavtseva, E. V. S., M.V.; Evstigneeva, R.P. "Some peculiarities of synthesis of cysteine-containing peptides." *Russ. Chem. Rev.,* **1998**, *67*, 545-562.

(30) Solomon, E. I.; Randall, D. W.; Glaser, T. "Electronic structures of active sites in electron transfer metalloproteins: contributions to reactivity." *Coord. Chem. Rev.,* **2000**, *200-202*, 595-632.

(31) Blasie, C. A.; Berg, J. M. "Kinetics and Thermodynamics of Copper(II) Binding to Apoazurin." *J. Am. Chem. Soc.,* **2003**, *125*, 6866-6867.

(32) Estrada, E.; Uriarte, E. "Folding degrees of azurins and pseudoazurins: Implications for structure and function." *Comp. Biol. and Chem.,* **2005**, *29*, 345-353.

(33) Gilardi, G. M., G.; Rosato, N.; Canters, G.W.; Finnazzi-Agro, A. "Unique Environment of Trp48 in Pseudomonas aeruginosa Azurin As Probed by Site-Directed Mutagenesis and Dynamic Fluorescence Spectroscopy." **1994**, *33*, 1425-1433.

(34) Malmstrom, B. G.; Wittung-Stafshede, P. "Effects of protein folding on metalloprotein redox-active sites: electron-transfer properties of blue and purple copper proteins." *Coord. Chem. Rev.,* **1999**, *185-186*, 127-140.

(35) Pozdnyakova, I.; Guidry, J.; Wittung-Stafshede, P. "Copper Stabilizes Azurin by Decreasing the Unfolding Rate." **2001**, *390*, 146-148.

(36) Pozdnyakova, I.; Wittung-Stafshede, P. "Copper Binding before Polypeptide Folding Speeds Up Formation of Active (Holo) Pseudomonas aeruginosa Azurin." *Biochemistry,* **2001**, *40*, 13728-13733.

(37) Tigerstrom, A.; Schwarz, F.; Karlsson, G.; Okvist, M.; Alvarez-Rua, C.; Maeder, D.; Robb, F. T.; Sjolin, L. "Effects of a Novel Disulfide Bond and Engineered

Electrostatic Interactions on the Thermostability of Azurin." *Biochemistry,* **2004**, *43*, 12563-12574.

(38) Wittung-Stafshede, P. "Role of Cofactors in Folding of the Blue-Copper Protein Azurin." *Inorg. Chem,* **2004**, *43*, 7926-7933.

(39) Johnson, E. C. B.; Kent, S. B. H. "Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction." *J. Am. Chem. Soc.,* **2006**, *128*, 6640-6646.

(40) Solomon, E. I.; Hedman, B.; Hodgson, K. O.; Dey, A.; Szilagyi, R. K. "Ligand K-edge X-ray absorption spectroscopy: covalency of ligand-metal bonds." *Coord. Chem. Rev.,* **2005**, *249*, 97-129.

(41) Solomon, E. I.; Penfield, K. W.; Gewirth, A. A.; Lowery, M. D.; Shadle, S. E.; Guckert, J. A.; LaCroix, L. B. "Electronic structure of the oxidized and reduced blue copper sites: contributions to the electron transfer pathway, reduction potential, and geometry." *Inorg. Chim. Acta.,* **1996**, *243*, 67-78.

(42) Solomon, E. I.; Randall, D. W.; Glaser, T. "Electronic structures of active sites in electron transfer metalloproteins: contributions to reactivity." *Coord. Chem. Rev.,* **2000**, *200-202*, 595-632.

(43) Levengood, M. R.; van der Donk, W. A. "Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies." *Nat. Protocols,* **2007**, *1*, 3001-3010.

CHAPTER 4:

TRANSFORMING BLUE COPPER TO RED COPPER: MET121HCY AZURIN[i,ii]

4.1 INTRODUCTION

Mononuclear copper sites are separated into two distinct classes, the type 1 or blue copper proteins and type 2 copper proteins. The blue copper proteins are characterized by their unique strong blue color due to an intense visible absorption around 600 nm ($\varepsilon \sim 5000\ M^{-1}cm^{-1}$).[1-3] They also exhibit narrow hyperfine coupling constants ($A_{||} < 100 \times 10^{-4}\ cm^{-1}$) in the electron paramagnetic resonance (EPR) spectrum and the spectra can be either axial ($g_z > g_y \approx g_x$) or rhombic ($g_z > g_y > g_x$) in nature.[1,2] The type 1 copper protein spectral features are in direct contrast to the type 2 copper proteins that have much weaker absorption in the visible region ($\varepsilon \sim 40\ M^{-1}cm^{-1}$) and wider hyperfine coupling constants ($A_{||} > 150 \times 10^{-4}\ cm^{-1}$).[4,5]

The success in defining the spectroscopic properties of each site has advanced our knowledge in not only biocoordination chemistry but also has elucidated the biological functions of copper proteins. The unique type 1 copper structure is responsible for the interesting spectroscopic features,[6-8] the high redox potentials,[9-13] and the efficient electron transfer (ET) rates seen in type 1 proteins.[14-18] As electron transfer proteins, type 1 copper proteins participate in a myriad of biological processes and are required to possess tunable redox potentials and ET rates to match those of their partners.[14-16,18] A current challenge in the field is to understand how type 1 copper proteins can tune their redox potentials and ET rates without losing their structural identity and functional

[i] This work has been done in collaboration with Dr. Ninian Blackburn's Laboratory at the Oregon Health and Science University. All EXAFS data, figures, and analysis were generated from runs at the Stanford Synchrotron Radiation Lightsource (Menlo Park, CA).
[ii] This work has been done in collaboration with Yang Yu and Nicholas Marshall from the Lu group at University of Illinois-Urbana.

properties. A major avenue for such tuning is through the interaction between the axial ligand and copper.[19,20] When the axial ligand is weakly coordinated to the copper ion, such as in azurin from *Pseudomonas aeruginosa* (Figure 4.1), the unique trigonal geometry makes it possible for the d_x^2-$_y^2$ ground state of Cu^{2+} to overlap most favorably with pπ of the Cys ligand, resulting in the short Cu-S bond and giving a strong pπ(Cys)-to-Cu^{2+} charge transfer band around 600 nm.[21,22] However, when the axial ligand becomes stronger, such as in nitrite reductase, a model called "coupled distortion" has been proposed where the strong axial ligand interaction with copper ion results in a correspondingly longer Cu-Cys bond and rotation of the d_x^2-$_y^2$ ground state of the Cu^{2+} ion so that it overlaps better with the pσ orbital of Cys.[22] Thus a more intense pσ(Cys)-to-Cu^{2+} charge transfer band at higher energy (~ 400 nm) is observed while the pπ(Cys)-to-Cu^{2+} charge transfer band around 600 nm becomes less intense. This change results in a strong green color instead of blue color, and therefore these proteins with stronger "coupled distortion" are often called green copper proteins. For different type 1 copper proteins, the relative absorption ratios of 400 nm and 600 nm is a good measure of how strong the axial ligand-copper and Cu-Cys bonds are.[23,24] The best manifest of the "coupled distortion" model is a recently discovered red copper protein called nitrosocyanin, that exhibits the strongest pσ(Cys)-to-Cu^{2+} charge transfer band around 390 nm and weakest pπ(Cys)-to-Cu^{2+} charge transfer band around 600 nm.[25-29] In addition, it also has a longer Cu-S(Cys) bond (2.26 Å) and wider $A_{\|}$ (150 x 10^{-4} cm^{-1}).[25-29]

Because of these interesting differences, a number of spectroscopic[6-8,30-37] and x-ray crystallographic studies[17,38-41] have been carried out to better define the structural origin of these spectral properties. Recent analysis using nuclear magnetic resonance

(NMR) spectroscopy has also helped to assign structural and electrostatic interactions.[42-44] Combined, these studies have established that the mononuclear copper bound by type 1 copper proteins is coordinated by one cysteine (Cys112 in azurin) and two histidines (His46 and His117 in azurin) in a trigonal plane. A weak axial ligand, often a methionine (Met121 in azurin) is present approximately perpendicular to the plane. The intense absorption at ~ 600 nm has been attributed to the allowed Cys ligand-to-copper charge transfer transition (LMCT) while the weak absorption bands in the same region for the normal type 2 copper proteins have been assigned as the parity-forbidden *d-d* transitions.[1-3] In addition, the narrow $A_{\|}$ and the EPR spectral shape have been ascribed to the short Cu-Cys bond found in type 1 copper sites (~ 2.1 Å).[1-3] This short bond results in strong covalent interaction between copper and Cys, allowing the unpaired electron on Cu(II) to be delocalized onto Cys. This delocalization makes the hyperfine coupling less effective than type 2 copper that has a normal Cu-Cys bond length (2.29 Å).[4,5]

4.2 Interconverting Blue And Green Copper Sites by Site Directed Mutagenesis

While it is interesting to show the "coupled distortion" model may explain the properties of different copper proteins, an ultimate test of the model and our understanding of the copper proteins is demonstration of the "coupled distortion" in a single protein by transforming one protein into another through axial ligand mutations. Replacement of the weak axial Met121 residue in azurin to stronger coordinating ligands such as His[45-48] and Glu,[13,38,49,50] resulted in proteins that displayed spectral similarities to green copper proteins, with narrow EPR hyperfine couplings ($A_{\|} < 100 \times 10^{-4}$ cm^{-1}). A strong pH dependence of the ratio between the pσ(Cys)-to-Cu^{2+} charge transfer band

around 400 nm and the pπ(Cys)-to-Cu^{2+} charge transfer band around 600 nm has been observed in these mutants.[38,45] Concurrently, a green copper protein nitrite reductase was converted into a blue copper protein through mutation of Met182 into the much weaker ligand Thr.[27]

However, examples in which a blue copper protein was transformed into a red copper protein have not been reported, probably because such a transformation requires replacing a weak axial ligand with a stronger ligand while maintaining a short bonding distance to the equatorial Cys112. Such a demonstration of a strong axial ligand interaction that transforms a blue copper into a red copper protein is an important piece of the puzzle in this "coupled distortion" model to explain both structural and functional tuning of this important class of ET proteins.

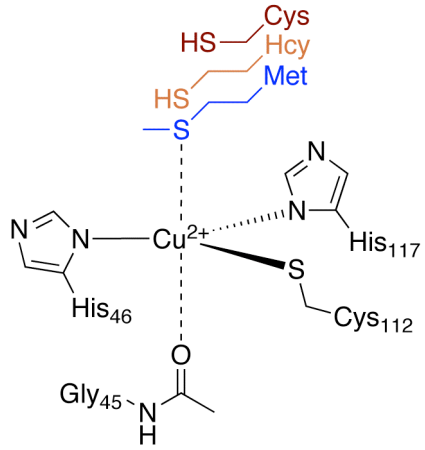


Figure 4.1: Schematic representation of the *Pseudomonas aeruginosa* azurin type I Cu site. The WT protein has an axial Met at residue 121. The two proposed substituted stronger ligands, Cys and Hcy, are shown in comparison to Met121.

4.3 Preparation of Met121Cys Azurin

Although azurin has been subjected to numerous site-directed mutagenesis studies to investigate the effects of substitution by proteinogenic and nonproteinogenic amino

acids on redox and ET properties, these studies have provided little information regarding how to create red copper sites in blue copper scaffolds to test the "coupled distortion" model.[19,51-54] Previous reports of Met121His and Met121Glu variants of azurin show the transformation of blue copper proteins into green copper proteins, but not into red copper proteins.[8,19,20,38,45,47,48,50-60] However, replacing the Met with a stronger ligand such as Cys, may provide the ligand strength necessary to create a red copper site and help confirm the "coupled distortion" model by possibly allowing more overlap of pσ(Cys) and the d_x^2-$_y^2$ ground state of Cu^{2+}. Although the Met121Cys mutation has been listed as part of Met121 site-saturated cassette mutagenesis studies, a full report characterizing an azurin containing a free thiolate amino acid in the Met121 axial position is not available.[13,52,53]

Therefore replacement of Met121 in azurin with cysteine was completed using standard site-directed mutagenesis protocols. Following verification of the correct nucleotide changes, the protein was expressed analogously to azurin mutants previously reported.[61] The yield of recombinant expressed Met121Cys azurin mutant was ~30 mg/L growth media, comparable to that of WT azurin. The protein was then directly titrated with $Cu(II)SO_4$ resulting in a color change from colorless to blue. UV-visible spectroscopy (Figure 4.3) verified a strong absorption band around 610 nm and a weaker band around 440 nm, reminiscent of the absorption spectra found for both the Met121His and Met121Glu proteins at low pH. ESI-MS and MALDI-MS analysis of the protein showed only a single copper ion bound in the protein.

4.4 UV-Vis Spectral Characterization of Met121Cys Azurin

A comparison between the UV-Vis spectrum of Met121Cys azurin with that of wild type (WT), Met121His and Met121Glu azurins reveals that the Met121Cys behaves like the Met121His and Met121Glu mutants. Like WT azurin, Met121Cys azurin displays a strong absorption band around 610 nm while displaying a weaker band at 420 nm (Figure 4.2, Table 4.1). Since the more intense peak is closer to the LMCT band found in WT azurin, we assigned the absorbance at 610 nm as the pπ(Cys)-to-Cu^{2+} charge transfer (CT) band, similar to how the bands were assigned for Met121His azurin and Met121Glu azurin at pH 4.[38,45] In contrast to WT azurin, but similar to Met121His and Met121Glu azurin at pH 4, the intensity of the absorbance at 440 nm, assigned to the pσ(Cys)-to-Cu^{2+} charge transferd is higher than that of WT azurin, suggesting that introduction of the stronger axial ligand Cys resulted in more overlap of pσ(Cys) and the $d_{x^2-y^2}$ ground state of Cu^{2+}. The ratio of the two peaks, RL [=ε(high energy)/ε(low energy)] can provide a means to measure contributions of each transition, as a lower RL value indicates more pπ(Cys)-to-Cu^{2+} charge transfer and more "blue" character while a higher RL suggests more pσ(Cys)-to-Cu^{2+} charge transfer and more "red" character.[18,24,57,62]

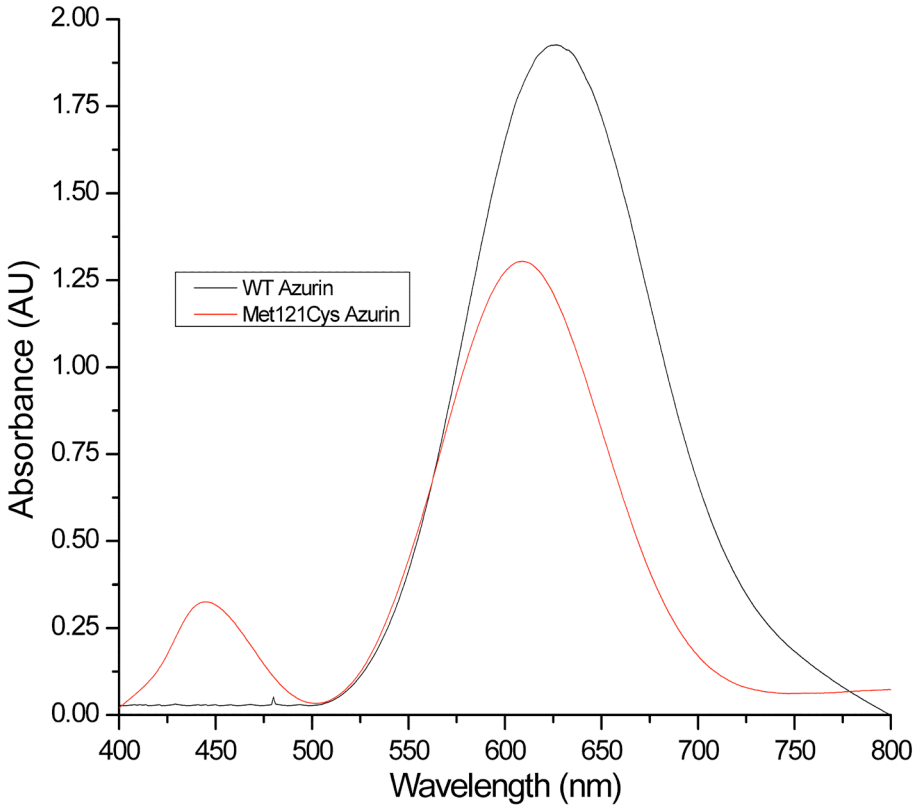


Figure 4.2: UV-visible spectrum for Met121Cys azurin at pH 5 in UB buffer (See experimental). Absorption bands are at 440 nm, and 610 nm.

While it is interesting to find that a stronger axial ligand (e.g., Cys) resulted in a higher ratio, a trend predicted by the "coupled distortion" model and found in both Met121His and Met121Glu azurin, the magnitude of the ratio is not consistent with the model, since the stronger Cys ligand should result in a much higher RL value than that of nitrite reductase because nitrite reductase has a Met at the axial ligand position. Therefore the Met121Cys copper site was subjected to a pH titration analysis to determine if at higher pH values, potential deprotonation of Cys121 would result in more interaction with the Cu(II) ion.

As shown in Figure 4.3, adjusting the pH values from 4.3 to 9.6 resulted in a decrease of the 610 nm band and increase of the 420 nm band, and thus the RL value.

Furthermore, the color of the protein changes from blue to blue-green as the pH is increased, following the predicted trend of "coupled distortion" associated with more pσ(Cys)-to-Cu^{2+} charge transfer. These changes are not found in similar experiments with WT azurin. Thus deprotonation of Cys121 at higher pH increases axial interaction with Cu. At high pH values (9.6), the RL ratio is one indicating equal contribution of CT from both Cys112 and Cys121. While suggestive of a stronger axial interaction, and in support of the blue-green "coupled distortion" theory, the strength of the axial interaction is too weak to attribute the changes to creation of a "red" Cu site. The changes occurring upon changing the pH are not due to a simple two-state equilibrium but likely involve multiple structural changes as isosbestic points are not observed.

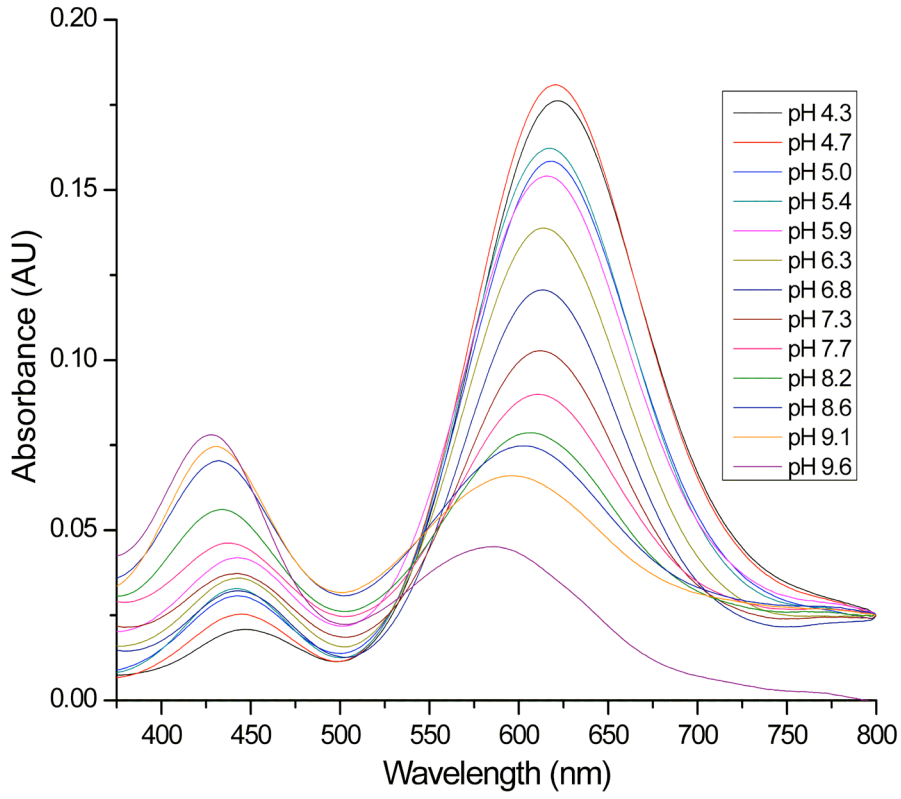


Figure 4.3: UV-visible spectrum for Met121Cys azurin at varying pH values. pH range is from 4.3-9.6 in UB Buffer.

4.5 X-BAND EPR CHARACTERIZATION OF MET121CYS AZURIN

X-band EPR spectroscopy was used to probe for changes in the site geometry (Figure 4.4). At pH 4, the observed Met121Cys azurin spectrum is very similar to that of WT azurin at pH 5. Though the hyperfine splitting is not as clear as that found in WT azurin, the hyperfine splitting of the mutant ($A_{\|}$ = 95 x 10^{-4} cm^{-1}) is similar to the type 1 copper site in WT azurin (Table 4.1). When the pH is increased, however, these splittings disappear with the generation of a new, more rhombic species with $A_{\|}$ = 195 x 10^{-4} cm^{-1}, characteristic of a type 2 copper-thiolate center. This site distortion is not found in WT azurin. These EPR spectral changes indicate that increasing pH may cause deprotonation of the axial Cys121 thiolate and thus stronger interaction with the Cu ion, which in turn resulted in significant change in the copper center geometry to a more type 2 copper.

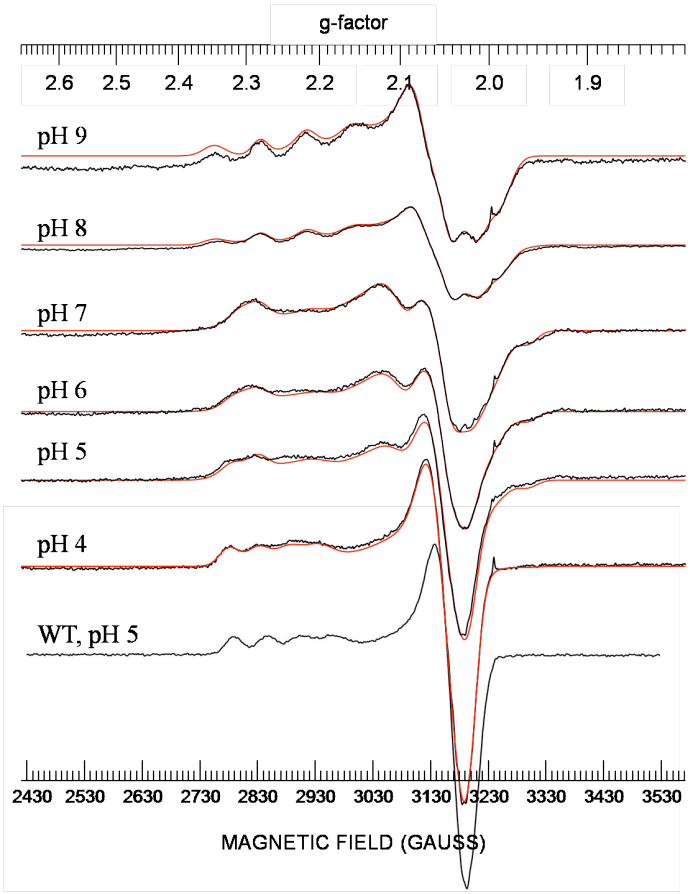


Figure 4.4: X-band EPR spectra of Met121Cys azurin at different pH values. The EPR spectra are in black while the simulated spectra are in red. X-band experimental conditions: microwave frequency, 9.04 GHz; power, 0.2 mW; modulation amplitude 5 G, time constant, 32 ms; sweep time, 60 s; number of scans, 10; gain, 8000; temperature, 29 K. The spike at *g*=2.0 signal is an artifact of the EPR tubes used.

Interestingly, the spectral changes monitored by UV-vis (Figure 4.3) do not correlate directly with spectral changes in the EPR spectra (Figure 4.4). For example, at pH 6, the UV-vis indicates minimal change in the RL from that at pH 5, while the EPR suggests significant changes from pH 5 to 6, with appearance of a type 2 copper species. The main difference between the conditions of the two spectroscopic methods is that the UV-Vis spectra were collected at room temperature while the EPR data were obtained at 29 K. It is well established that the pH of buffers is dependent on temperature[63] and the

discrepancies between the EPR and UV-Vis spectra seen here are likely attributable to pH changes in the buffering system upon cooling to 29 K that was required for EPR spectroscopy. We, therefore, tested the pH change upon cooling in the buffer system used for this study according to the previously published method and found that the increased rhombicity found in the EPR spectra was reduced, suggesting the EPR spectra are not reminiscent of the true pH values before freezing with the temperature independent buffer.[63] The reported EPR spectra of Met121Cys from pH 5-9 are, therefore, actually at higher pH than it is indicated at room temperature. This temperature effect on pH caused a higher proportion of deprotonated Cys and more distorted tetragonal, rhombic, character to the EPR spectra. At intermediate pH values, where mixed deprotonation states of the Cys ligands would exist, a mixture of the two species can be seen in the EPR spectrum (Figure 4.4).

Table 4.1: EPR parameters for Met121X variants.

	pH	λ max (nm)	R_L values ($^{\sim 400nm}/_{\sim 600nm}$)	$A_{\|\|}$ (10^{-4} cm^{-1})	$g_{\|\|}$
WT Azurin	5	628	NR	66	2.25
Met121His	7	440, 600	2	102	2.24
Met121Glu	7	450, 615	2.5	83	2.3
Met121Cys	4	451, 627	0.05	54*	2.26*
Met121Cys	5	447, 621	0.15	54*	2.26*
Met121Cys	6	444, 615	0.23	18*	2.29*
Met121Cys	7	441, 612	0.29	18*	2.29*
Met121Cys	8	434, 608	0.55	87*	2.25*
Met121Cys	9	430, 591	1.47	87*	2.25*
Met121Hcy	7	410, 590	1.5	180	2.25
Nitrosocyanin	7	390, 500	2.5	150	2.24

The presence of different species in the EPR spectrum at different pH for the Met121Cys azurin variant may be attributed to site destabilization. The anticipated

Cys121-Cu bond length (> 5 Å) is exceptionally long and lies outside the expected distance required for a strong Cys(S)-Cu covalent interaction. This distance is expected to result in a very weak axial Cys-Cu interaction. This interaction would be further weakened if the Cys121 thiol remains protonated. Based on the UV-Vis and EPR data, increased pH results in deprotonation, likely at Cys121 (similar spectral changes are not noted in WT Azurin suggesting Cys112 remains bound to the Cu), and seems to strengthen the Cys121-Cu interaction causing significant re-arrangement of the copper site. Thus because the Cys seems too short for the thiolate to fully interact with the Cu, incorporation of a nonproteinogenic amino acid with a longer side chain may allow an axial thiolate to interact with the Cu. Homocysteine (Hcy) was chosen as the Cys derivative because its side-chain is one methylene group longer than Cys and would therefore be in approximately the same position as the Met thioether would reside. The decreased distance should permit interaction with the copper at any pH, and result in a stable "coupled distorted" complex.

4.6 Preparation of Met121Hcy Azurin

Replacing Met121 with Hcy should provide a more stable copper center than Cys since the extra methylene group in Hcy can provide a shorter bond distance between the thiolate and copper. As discussed in chapter 2, EPL can be used to incorporate Hcy in place of Cys112. Furthermore, our laboratory has successfully shown that substitution of Met121 with nonproteinogenic amino acids can be accomplished using EPL (See Figure 1.15).[19,20] In the presence of synthetic peptides containing free thiols in the side chain of the N-terminal residue such as Cys, the protein thioester can undergo a chemoselective

transthioesterification to perform native chemical ligation creating a protein that is indistinguishable from recombinantly expressed WT protein (Figure 4.5).

Therefore, similar to studies discussed in chapters 2 and 3, the synthesis of a 17-mer peptide consisting of the final C-terminal 17 amino acids of azurin with a Met121Hcy mutation Met was completed. The typical yields of the purified peptide were approximately 60-80 mg of peptide for a 0.1 mmol synthesis affording ample material for semi-synthesis of Met121Hcy azurin using EPL. EPL reactions were first carried out in the presence of sodium 2-mercaptoethanesulfonate (MESNa) or N-methylmercaptoacetamide (NMA) as the transthioesterification catalyst. However, as discussed in chapter 3, recent work from the Kent group demonstrated that the nucleophilicity and leaving group ability of the catalyst play key roles in the rate of transthioesterification.[66] Topping the list with the fastest transfers and reduced hydrolysis rates was 4-mercaptophenylacetic acid (MPAA).[66] In this work, reactions with all mediators yielded correctly ligated full length (FL) mutant proteins but ligations with MPAA resulted in yields 2x-3x higher than the other mediator used. Thus all subsequent ligations were performed using MPAA. In subsequent attempts to further increase the yield of the Hcy protein variant, a modified procedure, prompted by the studies discussed in chapter 3, was performed in which an anaerobic environment was maintained for the duration of the ligation. This modification resulted in moderate yield improvements (improvement of 2-3 mg/L growth) over using aerobic conditions and was used thereafter.

4.7 UV-Vis Spectral Characterization of Met121Hcy Azurin

UV-Vis spectroscopy was used first to characterize Met121Hcy azurin. Upon addition of Cu(II) to apo-Met121Hcy azurin, a stronger peak was observed around 410 nm, together with a weaker peak around 590 nm (Figure 4.5). This higher RL ratio is in direct contrast to that of WT azurin and different from the Met121Cys azurin at low pH. If not coordinated to the copper, the Hcy thiolate is expected to remain protonated at low pH (as the calculated pKa value for Hcy is ~10) and resulting in spectra similar to Met121Cys azurin. However, the RL ratio is reversed from Met121Cys azurin, suggesting a strong interaction between the copper ion and Hcy. Most importantly, the spectrum is remarkably similar to that seen in the red copper protein nitrosocyanin even though the molar absorptivities of the peaks are lower (Figure 4.5, insert).[25-29]

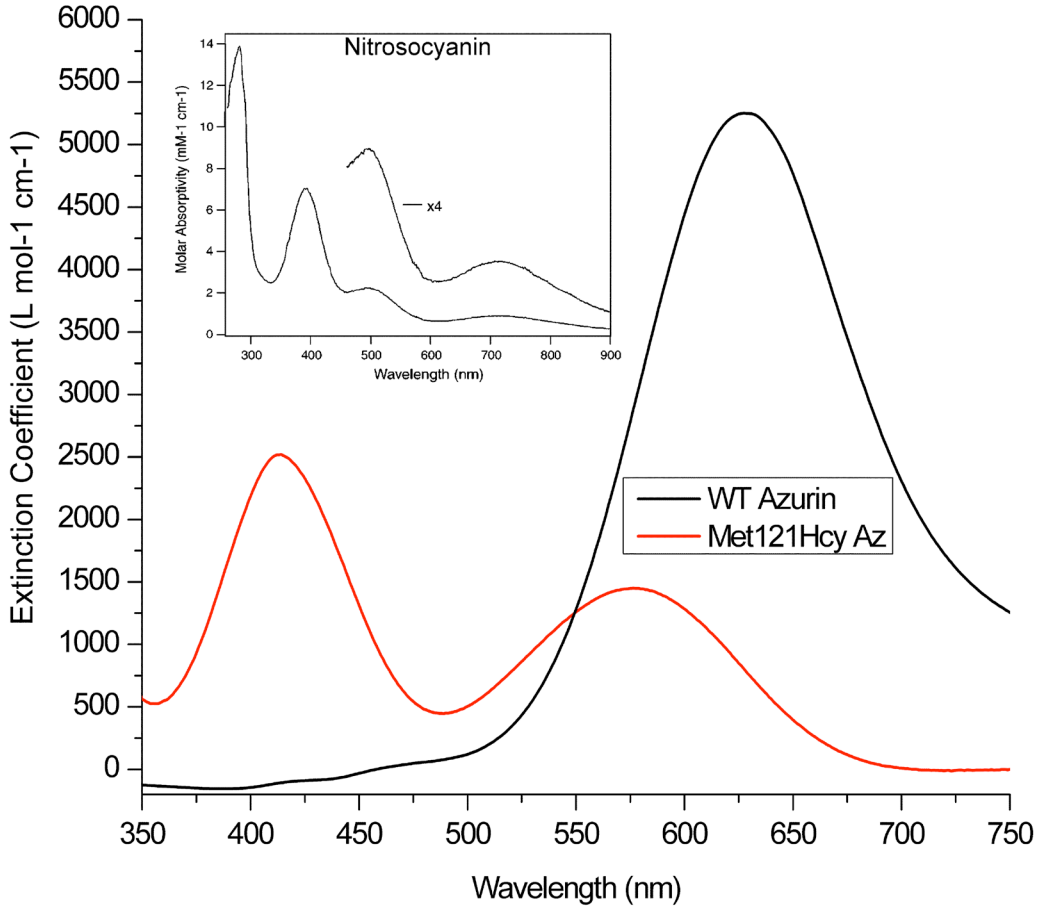


Figure 4.5: Visible spectra of WT and Met121Hcy azurin in 50 mM MOPS, pH 7. Nitrosocyanin (insert) shows absorptivities at 400 nm and 510 nm.[25]

The apo protein was refolded and titrated anaerobically using Cu(I) and Cu(II) to verify that both thiolates (Cys112 and Hcy121) were free to participate in Cu binding and are not involved in a disulfide. No significant changes were observable. Because binding Cu could potentially cause oxidation and create a disulfide between Hcy121 and Cys112, holo-protein samples were also titrated with Cu(II) and Cu(I) in the presence of tris(2-carboxyethyl)phosphine (TCEP). Again, no spectral changes were observable in the presence of TCEP, suggesting that the final holo-Met121Hcy azurin variant is fully reduced both prior and after Cu binding (Figure 4.6).

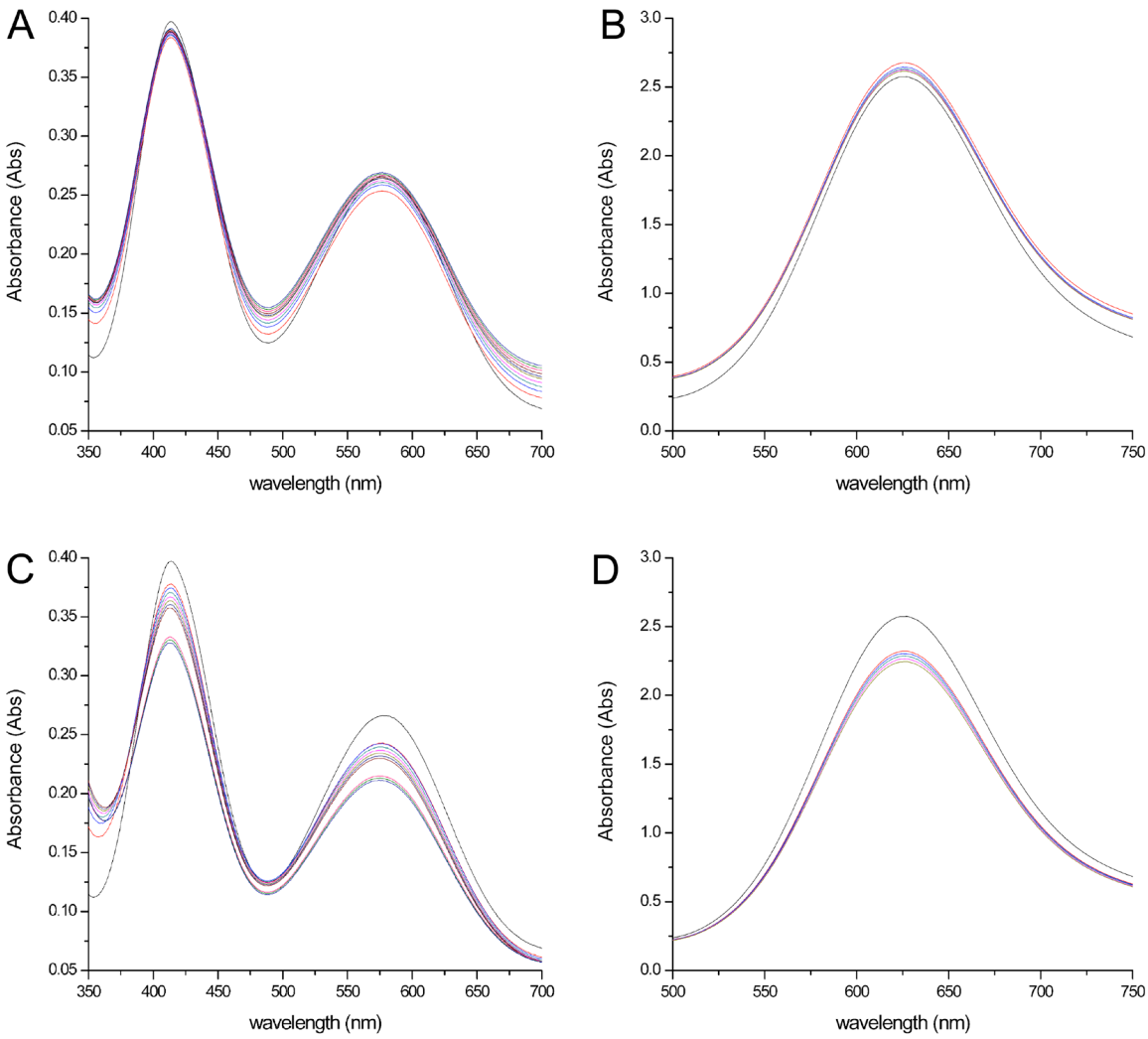


Figure 4.6: Met121Hcy titrations in the presence of TCEP. A) Met121Hcy titrated with 0.25 equivalents of TCEP. No changes to the LMCT peaks are noticeable. B) WT Azurin titrated with 0.25 equivalents TCEP. No changes to the LMCT peak are present. C) Met121Hcy azurin titrated with 1 equivalent TCEP. Only baseline shifts are noticeable and attributable to reduced protein that has been unfolded in the presence of TCEP. D) WT azurin titrated with 1 equivalent of TCEP. Similarly only baseline shifts are noticeable due to protein unfolding.

4.8 X-Band EPR Characterization of Met121Hcy Azurin

The similarity of Met121Hcy to red copper protein is further supported by EPR spectroscopy. The X-band EPR spectrum shown in Figure 4.8 displays a much larger parallel hyperfine splitting in the Met121Hcy mutant when compared to both WT azurin,

the green copper azurin mutants (Met121His and Met121Glu) as well as the Met121Cys mutant at low pH (Table 4.1). Both the shape and the hyperfine splitting of the spectrum are almost identical to that of red copper nitrosocyanin (Figure 4.7, Table 4.1).[26-29] The additional hyperfine splitting along $g_{\perp}$ indicates a stronger interaction between the copper and two equivalent nitrogen atoms, a feature also found in native nitrosocyanin.[25,26] The A values of ~15 x 10^{-4} cm^{-1} for the hyperfine pattern are also consistent with nitrogen ligands coordinated to a copper.[25,26] In contrast to the EPR spectrum for Met121Cys, the EPR spectrum for Met121Hcy indicates a single species at multiple pH values, suggesting that the M121Hcy azurin is a more stable site and not subject to pH dependent changes (Figure 4.7). Surprisingly, in contrast to Met121Cys azurin, Met121Hcy azurin is unstable and looses copper at pH values lower than 5 and higher than 8 (data not shown).

These results strongly suggest that introducing the nonproteinogenic amino acid Hcy has transformed the blue copper protein, Az, into a red copper protein; the extra methylene group provides sufficient length to allow for a strong Hcy-Cu interaction. Previous demonstrations of the "coupled distortion" model have focused on replacing a medium strength ligand such as Met with weaker ligands, such as Thr.[27] This is the first time that a medium strength ligand has been replaced with a much stronger ligand (thioether vs. thiolate). Though Met121Hcy shares spectroscopic similarities with nitrosocyanin, it was imperative to identify the ligand set bound to the Cu ion. The nitrosocyanin ligand set does not contain two thiolates, rather an exogenous water is bound fulfilling the Cu coordination sphere. Therefore it was necessary to perform EXAFS to eliminate any possibilities of exogenous ligands such as water.

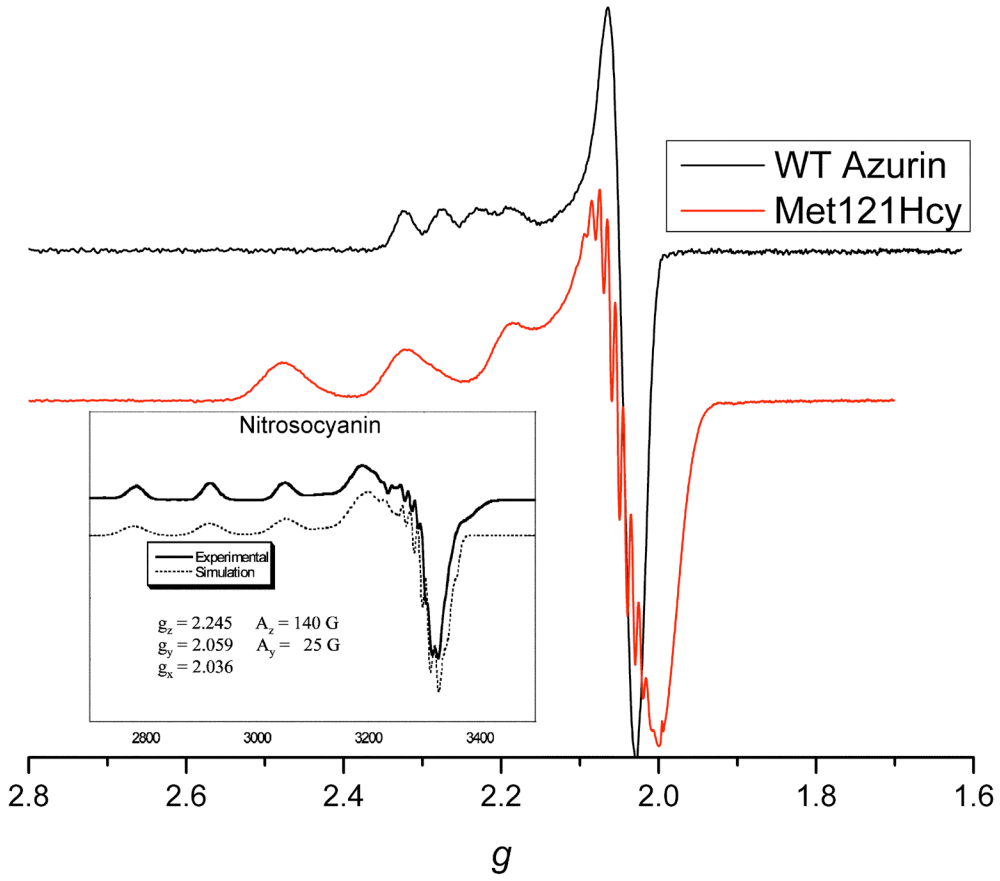


Figure 4.7: Met121Hcy azurin and WT X-band EPR spectra. X-band experimental conditions: microwave frequency, 9.05095 GHz; power, 0.2 mW; modulation amplitude 5 G, time constant, 32 ms; sweep time, 60 s; number of scans, 30; gain, 1000; temperature, 29 K. Nitrosocyanin EPR spectrum is shown in the insert for comparison.[26]

4.9 EXAFS Characterization of Met121Cys and Met121Hcy Azurins

In order to determine the coordination environment of Cu(II) in the Met121Cys and Met121Hcy derivatives in more details, EXAFS data was collected by Dr. Ninian Blackburn, calibrating the data against the well studied EXAFS-derived structure of the WT azurin.[19] EXAFS data for the WT and Met121 variant proteins were generated by Dr. Blackburn and are shown in Figure 4.10 with the EXAFS generated structural parameters listed in Table 4.2. As expected, the WT protein can be simulated with two histidines and

one cysteine residue coordinated to Cu at distances of 1.95 and 2.18 Å respectively (Figure 4.8a). The latter distance is 0.04 Å longer than reported previously, but is within the range of values typically reported by EXAFS and crystallography for the Cu(II)-Cys112 interaction.[30,67]

For the Met121Cys variants the data show definitively that only a single cysteine coordinates as a strongly bound (covalent) ligand with a Cu-S bond length of 2.17 Å. Therefore we investigated whether the thiolate at residue 121 could form a longer weaker bond to copper. At pH 5, a slight improvement (13%) in the goodness of fit parameter was obtained when an additional S was included in the fit refining to 2.76 Å, while at pH 7 the data is unambiguous in defining the presence of an additional Cu-S interaction at 2.74 Å (Figure 4.8b). Here the improvement in the fit is dramatic when the longer distance to a second S is included with the F value dropping by 55%. Figure 4.8c shows the result of eliminating the long distance sulfur ligand, which causes inadequate intensity in the Fourier transform (FT) peak around R=2.7 Å.

Since the pH values were recorded at room temperature and the effective pH of the buffer used for these experiments increased with decreasing temperature, the actual pH values at 29 K should be higher. In addition, the UV-vis and EPR data suggest mixture of species at both pH 5 and 7. Therefore these results suggested the presence of mostly mono-thiolate copper interaction at low pH, with the second Cys, most likely Cys121 remaining protonated and not interacting with copper strongly. At high pH, however, deprotonation of a second thiolate, likely Cys121, resulted in coordination of two thiolates to Cu(II) with the second Cys121(S)-(Cu) bond at a distance equivalent to that of the native Met121. This suggests that protein-folding constraints prevent the

thiolate from bonding more strongly to the copper center and causes the geometric strain noted in the EPR spectra. This contrasts with other $(Cys)_2His$ ligand sets such as that found in Cu(II)-Sco1 where both Cys residues coordinate Cu(II) at an average distance of 2.21 Å.[68]

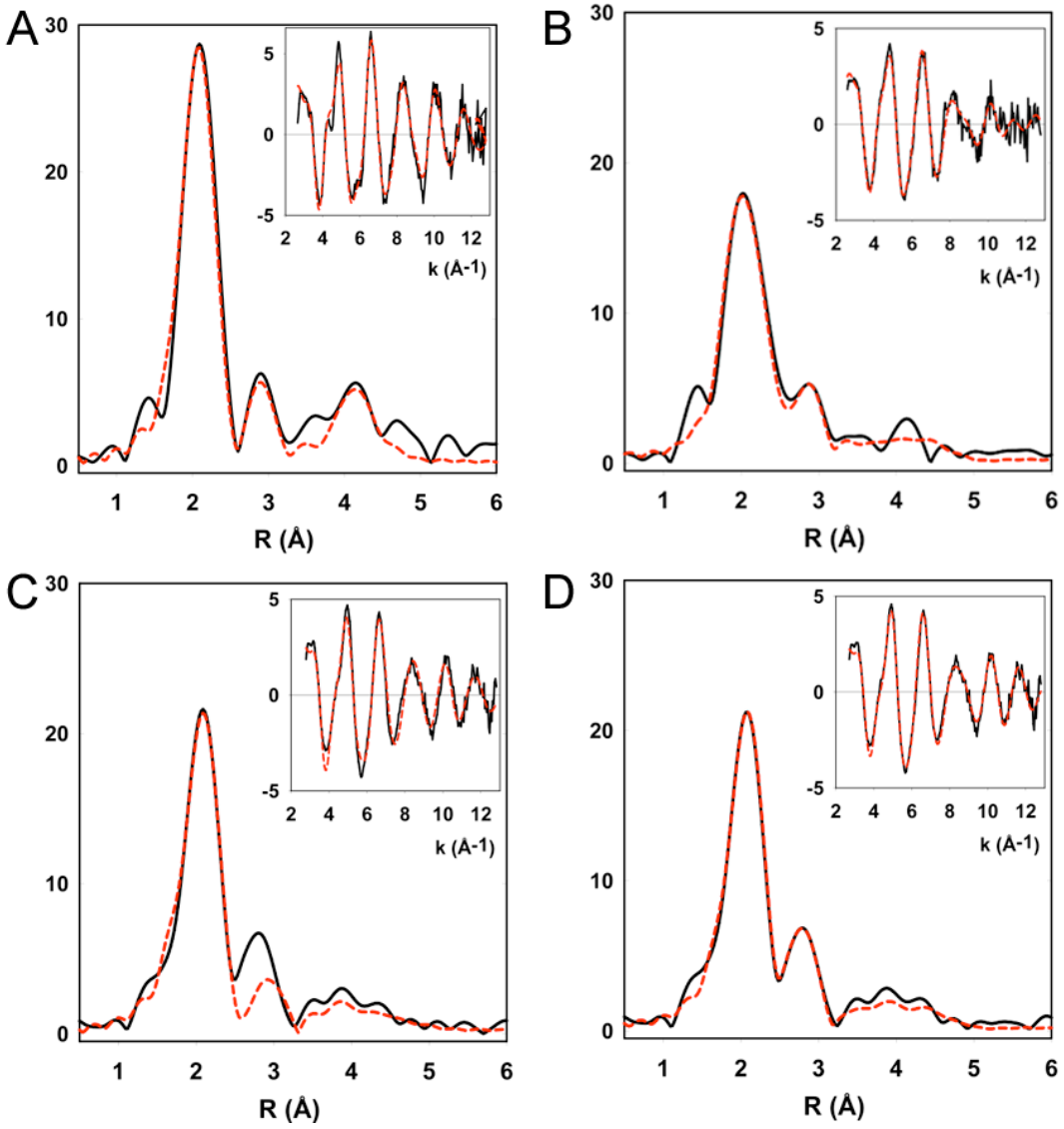


Figure 4.8: Fourier transform and EXAFS (inset) for WT azurin, Met121Cys and Met121Hcy variants. Experimental data are shown as solid black lines and the simulation is shown by dashed red lines. Parameters used to fit the data are listed in Table 4.2. A) WT azurin at pH 5.0; B) Best fit of Met121Cys variant at pH 7 (bolded in Table 3) with one short (2.17 Å) and one long (2.74 Å) Cu-S interaction; C) The best fit of Met121Cys variant at pH 7, obtainable using only the short Cu-S distance; D) Met121Hcy variant at pH 7. Spectra were collected and analyzed by Dr. Ninian Blackburn.

Table 4.2: EXAFS fitting parameters. Best fits are shown in bold.

Sample/ Fit	F[a]	Cu-S			Cu-N(His)[b]			E_0
		No[c]	R (Å)[d]	DW ($Å^2$)[e]	No[c]	R (Å)[d]	DW ($Å^2$)[e]	
WT								
A	**0.402**	**1**	**2.18**	**0.003**	**2**	**1.95**	**0.009**	**-0.763**
Met121Cys pH 5								
A	0.285	1	2.17	0.009	2	1.96	0.015	0.486
		1	2.76	0.014				
B	0.326	1	2.17	0.007	2	1.93	0.014	0.326
Met121Cys pH 7								
A	**0.165**	**1**	**2.17**	**0.008**	**2**	**1.95**	**0.02**	**0.37**
		1	**2.74**	**0.012**				
B	0.366	1	2.16	0.007	2	1.93	0.02	1.846
Met121Hcy pH 7								
A	**0.388**	**1**	**2.22**	**0.014**	**2**	**1.97**	**0.012**	**-2.05**
		1	**2.79**	**0.03**				
B	0.486	1	2.2	0.014	2	1.96	0.012	-2

[a] F is a least-squares fitting parameter defined as $F^2 = \frac{1}{N}\sum_{i=1}^{N} k^6 (Data - Model)^2$

[b] Fits modeled histidine coordination by an imidazole ring, which included single and multiple scattering contributions from the second shell (C2/C5) and third shell (C3/N4) atoms respectively. The Cu-N-C_x angles were as follows: Cu-N-C2 126°, Cu-N-C3 -126°, Cu-N-N4 163°, Cu-N-C5 -163°.

[c] Coordination numbers are generally considered accurate to ± 25%

[d] In any one fit, the statistical error in bond-lengths is ±0.005 Å. However, when errors due to imperfect background subtraction, phase-shift calculations, and noise in the data are compounded, the actual error is probably closer to ±0.02 Å.

[e] Debye-Waller factors are quoted as $2\sigma^2$

The best fit for the EXAFS spectrum of Met121Hcy Az shows two Cu-S interactions (Figure 4.8d, Table 4.2, bolded), presumably arising from Cys112 and Hcy121. The shorter Cu-S interaction in Met121Hcy, which is likely from Cys112, is the same as the Cys(S)-Cu bond length in WT Az. The longer of the two Cu-S bonds is similar in length, but slightly shorter than the Met(S)-Cu distance in Wt Az (2.77 Å vs.

2.9 Å). The shortening of this interaction is likely a consequence of the thiolate of Hcy121 being a stronger copper ligand when compared to the thioether in Met.

Therefore, the EXAFS spectra, in conjunction with the EPR and UV-vis data, support the model that Hcy is sufficiently long to form a strong copper-thiolate interaction regardless of temperature or pH, resulting in a "coupled-distorted" cupredoxin site. Cys at position 121 in Az is, conversely, too short to strongly interact with the copper and remains as a thiol at low pH. If Cys121 is deprotonated by raising the pH, the resulting thiolate becomes a copper ligand and distorts the copper site to a Type 2 copper site.

4.10 ELECTROCHEMICAL STUDIES

Other than the Met121Cys and Met121Hcy mutations characterized here, no other free thiolate mutation in the axial position 121 has been described. Our successful generation of these proteins allows us to address the question as to how the reduction potentials of a bis-thiolate site in azurin have changed in comparison to WT azurin. To probe the effects of a stronger ligand on the redox behavior of azurin, the redox potentials of Met121Cys and Met121Hcy azurin were measured using cyclic voltammetry. In comparison with WT azurin and the "green" azurin mutants (Met121His and Met121Glu), replacing Met121 with Cys resulted in lowering the reduction potential from 345 mV to 246 mV while replacing Met121 with Hcy resulted in lowering the potential to 295 mV at pH 5.0 (Table 4.3).[13,38,46,48,60] Since the reported reduction potential of nitrosocyanin[25] (~80 mV vs. NHE) was measured at pH 7.0, we repeated the reduction potential measurement of our Met121Cys and Met121Hcy azurins at the same

pH. Interestingly, at pH 7.0, the Met121Cys azurin has an almost identical reduction potential (83 mV) to that of nitrosocyanin, while the Met121Hcy azurin has a slightly higher potential of 119 mV, even though Met121Hcy azurin is much closer in spectroscopic properties to nitrosocyanin than Met121Cys azurin. Furthermore the Met121His ($E_{1/2}$ = 350 mV) or Met121Glu ($E_{1/2}$ = 370 mV) azurins have reduction potentials that are ~200 mV higher than that of nitrosocyanin at pH 7.[13,38,46,48,60] According to "coupled distortion" model, the stronger axial ligand should result in more tetragonal geometry thus lowering the reduction potential. Though WT azurin and the "green" azurins have been shown to display a decrease of their reduction potentials with increased pH, this trend is more noticeable in the mutants described in this chapter.

Table 4.3: Redox value comparison for azurin variants and nitrosocyanin. The comparison is between WT azurin, "green" azurins (Met121His/Glu), Met121Cys/Hcy and Nitrosocyanin at pH 5 and pH 7.

Mutant	$E_{1/2}$ pH 5 (mV)	$E_{1/2}$ pH 7 (mV)
WT Azurin	345	222
Met121His	~350	200
Met121Glu	~370	220
Met121Cys	264	83
Met121Hcy	295	119
Nitrosocyanin	NR	85

The Met121Cys/Hcy mutants display reduction potentials very similar to nitrosocyanin. Comparison of the nitrosocyanin site to azurin reveals that the dipole of azurin runs perpendicular to the equatorial plane and through the axial Met ligand at a slight angle.[26] In contrast, in nitrosocyanin the dipole axis travels through the equatorial, stronger ligands.[3,26,36] Because the spectroscopic properties of Met121Hcy are so similar to nitrosocyanin it suggests that the geometry of the site has significantly been altered so

that the 121 position may no longer be axial to the Cu in the Met121Hcy azurin variant, instead possibly making one of the His ligands axial and the new equatorial plane consisting of Hcy121, Cys112, and the remaining His. This helps support the larger parallel hyperfine splitting and tetragonal geometry seen in Met121Hcy azurin and not in Met121Cys azurin. Additionally, Met121Hcy azurin displayed scan rate dependence in CV, that is also not a feature in Met121Cys, suggesting perturbed electron transfer kinetics and a change in electronic structure.

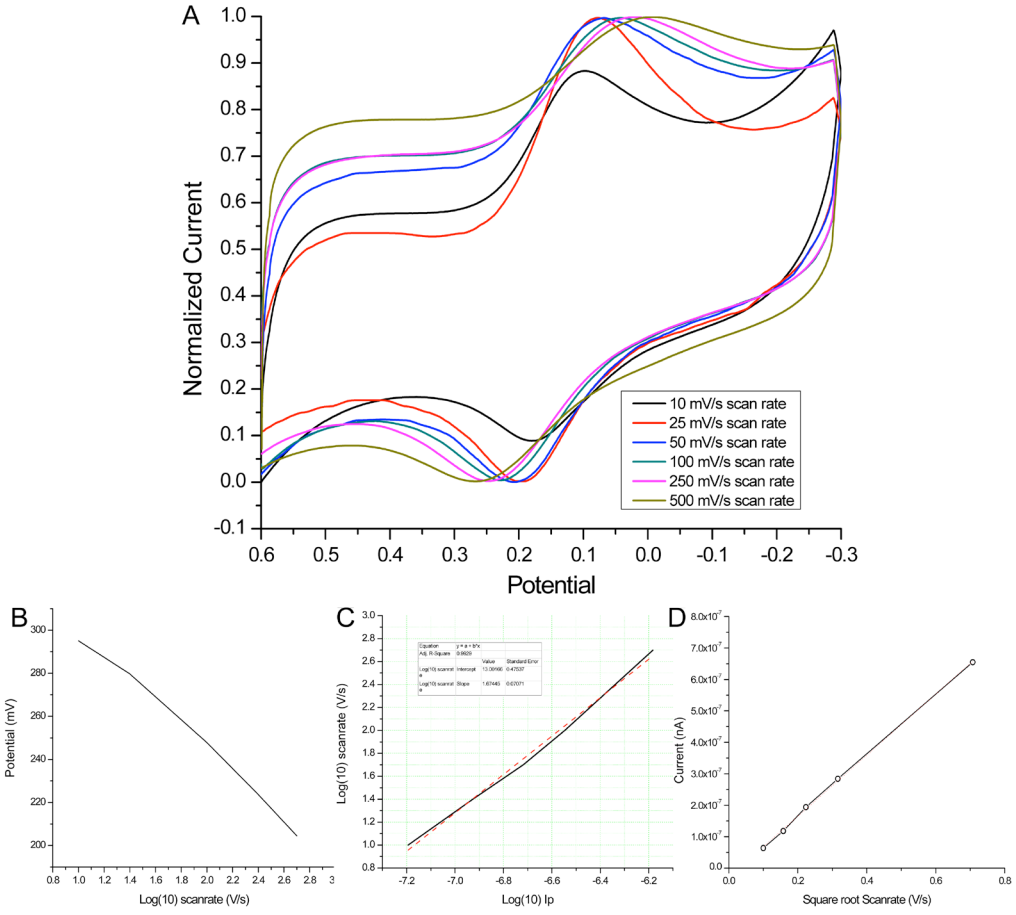


Figure 4.9: Cyclic voltammetry for Met121Hcy azurin. A) CV scans show strong scan rate dependence. As the scan rate is increased the potential decreases. Graphs B-D all demonstrate the dependency of the redox event on the scan-rate: B) Scan-rate vs. Ip C) Potential vs Scan-rate D) Current vs. Scan-rate.

4.11 Conclusions

In conclusion, successful incorporation of thiolate functional groups as strong axial ligands into azurin has been performed. Replacing Met121 with Cys using site-directed mutagenesis resulted in more tetragonal character, but not enough distortion to transform blue azurin to a red copper spectral mimic, due to longer and thus weaker interaction between Cys and Cu. Interestingly, adding an extra methylene group by introducing unnatural amino acid Hcy in azurin at the same position using EPL resulted in much stronger interaction and converted the blue site to a red copper site completely. The results firmly establish the "coupled distortion" model and demonstrate the power of using nonproteinogenic amino acids to address critical chemical biological questions.

4.12 Future Directions

To fully compare the engineered red copper site with nitrosocyanin, a crystal structure would be useful. While spectroscopically, the two proteins appear very similar it would be useful to see the changes in the geometry of the site and make a direct comparison of those changes with the WT nitrosocyanin site. To fully determine the electronic changes of the Met121Hcy site and compare it to the site in nitrosocyanin ESEEM and XANES are needed to determine where the dipole axis lies and to see if the Hcy121 still resides in the axial position.

4.13 Experimental

General. All chemicals were purchased from Sigma-Aldrich and were reagent grade or higher purity, and used without further purification.

Peptide Synthesis. All amino acids were purchased from Chem-Impex International (Wooddale, IL) and Advanced Chemtech (Louisville, KY). Protecting groups used for Fmoc-SPPS were Cys(Trt), Ser(OtBu), Glu(OtBu), Cys(StBu), Asn(Trt), His(Trt), Lys(Boc), Thr(OtBu). Preloaded Fmoc-Lys(Boc)-Wang resin was purchased from Advanced Chemtech and used after swelling for 3 x 8 min in dimethylformamide (DMF). Methylene chloride (DCM) was distilled from calcium hydride. Tetrahydrofuran (THF) was distilled from sodium and benzophenone. All aqueous buffers were degassed under Ar or N_2 as specified immediately prior to use.

UV-Vis spectra were taken on an HP Diode Array Spectrometer or a Cary 5000 spectrometer. X-band EPR spectra were collected on a Varian E-122 spectrometer at the Illinois EPR Research Center (IERC). The samples were run at ~30 K using liquid He and an Air Products Helitran cryostat without glassing agents. The magnetic field was calibrated using a Metrolab PT 2025 NMR Gaussmeter, and the microwave frequency was measured with an EIP model 578-frequency counter equipped with a high-frequency option.

Mass spectral data was collected by the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois by either Electrospray Ionization (ESI MS), or matrix assisted laser desorption ionization (MALDI) techniques as indicated. Mass spectral data reported as – mass, peak (percent).

Preparative RP-HPLC was performed with a Waters Delta 600 system. Solution A was 0.1 % TFA in H_2O, and solution B was 80 % MeCN/20 % H_2O with 0.1 % TFA. Unless otherwise stated, a linear gradient of 20% to 70% B over 23 min was used for all

runs. Solid phase peptide synthesis (SPPS) was performed on a Rainin PS3 automated peptide synthesizer or manually using medium fritted funnels.

Met121Cys Azurin Preparation and Expression, Lab Book 5, page 199.

A construct containing a methionine-to-cysteine mutation at position 121 in azurin was created using the Quikchange mutagenesis method (Stratagene, La Jolla, CA). The gene for WT azurin located in a pET9a plasmid was used as the template. The primers used were as follows:

-Forward: 5′-CACTCCGCACTGTGCAAAGGTACCCTG-3′
-Reverse: 5′-CAGGGTACCTTTGCACAGTGCGGAGTG-3′

The parental DNA was digested with Dpn1 at 37 °C for 2 h following the completion of PCR and *E. coli* BL21 (DE3) cells were transformed with the plasmid. The mutation was confirmed by sequencing of the insert DNA. The protein was then over-expressed and purified using methods analogous to previous protocols. The protein was titrated with $Cu(II)SO_4$ as a single batch and exchanged into a Universal Buffer (UB) buffer (50 mM NaOAc, 40 mM each MES/MOPS/TRIS/CAPS) at pH 5.0 by G25 Sephadex PD10 columns (GE Lifesciences).

Met121Cys Cu Titrations At Varying pH, Lab Book 7, page 55.

Holo-Met121Cys azurin protein was exchanged into UB buffer at the desired pH values. Samples were immediately placed in quartz cuvettes and a spectrum taken upon reaching the desired pH value.

General Procedure for Refolding EPL Products, Lab Book 7, page 37.

Dithiothreitol (DTT) (final concentration = 0.9 mM) and guanidinium hydrochloride (GdmHCl) (final concentration = 4 M) were added directly to a concentrated protein sample (~ 1 mL). After full dissolution, the sample was then left to sit at 25 ºC without disruption for 20 min. The denatured sample was then added dropwise to a 50-fold excess volume of 50 mM UB buffer, pH 5 at 4 ºC containing DTT (0.9 mM) at 1 mL/min at stirred for 10 min. The refolded sample was then immediately exchanged into fresh 50 mM UB buffer, pH 7.0 at 4 ºC using G25 Sephadex. Following concentration of the sample, the protein was titrated with either $Cu(II)SO_4$ (10 mM) or tetrakis(acetonitrile)copper(I) hexafluorophosphate (10 mM).

General Procedure for Titrating Proteins, Lab Book 7, page 37.

Samples titrated with Cu(II) were titrated with subequivalent additions of Cu(II) until saturation of the LMCT at 590 nm. Samples titrated with Cu(I) were titrated with subequivalent additions of Cu(I) to the previously established saturation point of the Cu(II) titrated sample. The samples that were titrated with Cu(I) were then exposed to gentle O_2 sparging overnight. The individually titrated samples were purified separately using a Pharmacia Q HiTrap HP column (5 mL) which had been washed with 10 column volumes of 1 M NaCl in H_2O and then equilibrated with 10 column volumes of 50 mM UB buffer, pH 7.

H_2N-CTFPGHSALMHcyKGTLTLK-COOH

Synthesis of Met121Hcy Azurin 17-mer Peptide (3.1). Lab Book 7, page 35.

Solid Phase Peptide Synthesis. The 17-mer peptide was synthesized on a 0.1 mmol scale with a Rainin model PS3 peptide synthesizer (Woburn, MA) using standard Fmoc-based chemistry. Synthesis from Lys128 through Lys122 was completed using a four-fold excess of amino acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and O-benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate (HBTU) (4 equiv) for 2 min prior to coupling to Wang resin preloaded with Fmoc-Lys(Boc) for 1-2 h. All Fmoc deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-Hcy(Trt)OH (3 equiv) was coupled to Lys122 in the presence of 1-hydroxybenzotriazole hydrate (HOBt) (6 equiv) and N,N'-diisopropylcarbodiimide (DIC) (6 equiv) in DMF at 25 °C for 2-4 h until full coupling could be observed by qualitative Kaiser tests. Synthesis was continued from Leu120 through Thr113 using a four-fold excess of amino acids pre-activated with 0.4 M NMM and HBTU (4 equiv) for 1-2 h. The N-terminal Cys112 was coupled as Boc-Cys(Trt)-OH (4 equiv) in the presence of HOBt/DIC (8 equiv each) with N_2 sparging for 5 h. Amino acids were double coupled when necessary and couplings were monitored by qualitative ninhydrin test.

*Peptide Cleavage***.** The resin carrying the fully protected peptide was washed with dichloromethane (3 x 15 mL) and dried under reduced pressure. Peptide cleavage from the resin was achieved by treating the resin with 2% triisopropylsilane (TIPS)/dH_2O/ethanedithiol (EDT) in neat trifluoroacetic acid (TFA) (5 mL) at 25 °C for 2 h. The cleaved peptide was isolated from the resin by filtration and the residual TFA was removed by rotary evaporation to afford a yellow oil. The peptide was precipitated by

trituration with cold diethyl ether and centrifuged at 2000 x *g* for 5 min. The peptide was iteratively washed with cold diethyl ether and centrifuged. The crude peptide was dissolved in 1:1 dH_2O/acetonitrile (~10-15 mL) and lyophilized prior to purification. The crude lyophilized product was purified by preparative RP-HPLC on a Waters Delta Pak C18 column (130 mm x 25 mm). The purified peptide was lyophilized directly after purification to afford a white solid and analyzed by MALDI-TOF-MS. M121Hcy Azurin 17-mer (H_2N-CTFPGHSALHcyKGTLTLK-OH), Expected: 1791.14, Observed: 1791.5.

Procedure for Expressed Protein Ligation, Lab Book 7, page 33.

Cultures of *E. coli* BL21 (DE3) cells containing a plasmid to express the azurin(1-111)-Intein-CBD fusion protein were grown in LB media for 8 h at 37 ºC and were used to inoculate eight 2 L flasks of LB media containing 100 mg/mL ampicillin. The cells were grown at 37 ºC for 16 h with shaking at 210 rpm. Protein expression was induced at ~16 h with 0.5 mM IPTG and induction was continued for 4 h at 37 ºC. The cells were then harvested at 9800 x *g*, stored at -20 ºC and used when needed.

Frozen azurin-thioester cell stock was re-suspended in a lysis buffer containing 25 mM HEPES, pH 7.2, 250 mM NaCl, 1 mM EDTA, 1 mM PMSF 0.1% Triton-X-100. The suspension was lysed using sonication (Misonix Sonicator 4000, 0.5 inch diameter probe) for a total work time of 9 min (6 sec. on, 10 sec. rest). The lysate was centrifuged at 13250 x *g* for 30 min. at 4 °C. The supernatant was carefully decanted. The fusion protein was then bound by batch absorption to 100 mL of chitin resin that had been pre-equilibrated with 20 mM HEPES, pH 8.0, 250 mM NaCl, and 1 mM EDTA (buffer 1) for 1-2 h at 4 ºC. The chitin resin was poured into a column and the column headspace was

purged with Ar. The column was then washed with 3 column volumes of buffer 1 under Ar pressure by cannulation.

Ligation was initiated by the addition of the 17-mer peptide (1.5 mM, 80 mg) and tris-(2-carboxyethyl)phosphine (TCEP) (1.5 mM, 14 mg) in 30 mL of buffer 1 containing 50 mM 4-mercaptophenylacetic acid (MPAA) under Ar and transferred directly to the column by Ar pressure. The chitin resin was then re-suspended in the column and the entire column was agitated gently at 4 ºC for 66 h.

After ligation, the column was eluted under Ar pressure and washed with 1 column volume of buffer 1. The eluent was centrifuged at 13250 x g for 30 min and the supernatant was concentrated using 10,000 MWCO Centricon concentration spin tubes at 2000 x *g* to a final volume of ~20 mL. The concentrated protein was then exchanged into 50 mM MOPS buffer, pH 7 using a G25 sephadex column.

Electrochemical Measurements, Lab Book 6, page 41.

Before running mutant samples, the WT azurin reduction potential was measure to standardize the electrode. Once stable, the reduction potential of each mutant (Met121Hcy and Mey121Cys Azurins) was determined by cyclic voltammetry using a CH Instruments 617A potentiostat equipped with a picoamp booster and a faraday cage. A pyrolytic graphite edge (PGE) electrode was polished and 2-3 mL protein solution was applied directly to the electrode following previously described methods. After a short incubation time, the electrode was immersed in either 20 mM NaOAc, pH 4.0 with 100 mM NaCl, 20 mM NH4OAc, pH 5.0 with 100 mM NaCl, or 25 mM KH_2PO_3, 100 mM KCl before data collection. Each protein was then sampled at varying scan rates between

10 mV and 500 mV. The reduction potentials were measured against Ag/AgCl and converted to NHE.

4.14 References

(1) Malmström, B. G. "Rack-Induced Bonding In Blue-Copper Proteins." *Eur. J. Biochem.,* **1994**, *223*, 711-718.

(2) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Roos, B. O. "Theoretical Study of the Electronic Spectrum of Plastocyanin." *J. Am. Chem. Soc.,* **1997**, *119*, 218-226.

(3) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. "Electronic Structures of Active Sites in Copper Proteins: Contributions to Reactivity." *Chem. Rev.,* **1992**, *92*, 521-542.

(4) den Blaauwen, T.; Hoitink, C. W. G.; Canters, G. W.; Han, J.; Loehr, T. M.; Sanders-Loehr, J. "Resonance Raman spectroscopy of the azurin His117Gly mutant. Interconversion of type 1 and type 2 copper sites through exogenous ligands." *Biochemistry.,* **1993**, *32*, 12455-12464.

(5) Han, J.; Loehr, T. M.; Lu, Y.; Valentine, J. S.; Averill, B. A.; Sanders-Loehr, J. "Resonance Raman excitation profiles indicate multiple Cys-Cu charge transfer transitions in type 1 copper proteins." *J. Am. Chem. Soc.,* **1993**, *115*, 4256-4263.

(6) Andrew, C. R.; Yeom, H.; Valentine, J. S.; Karlsson, B. G.; van Pouderoyen, G.; Canters, G. W.; Loehr, T. M.; Sanders-Loehr, J.; Bonander, N. "Raman Spectroscopy as an Indicator of Cu-S Bond Length in Type 1 and Type 2 Copper Cysteinate Proteins." *J. Am. Chem. Soc.,* **1994**, *116*, 11489-98.

(7) DeBeer George, S.; Basumallick, L.; Szilagyi, R. K.; Randall, D. W.; Hill, M. G.; Nersissian, A. M.; Valentine, J. S.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. "Spectroscopic Investigation of Stellacyanin Mutants: Axial Ligand Interactions at the Blue Copper Site." *J. Am. Chem. Soc.,* **2003**, *125*, 11314-11328.

(8) den Blaauwen, T.; Canters, G. W. "Creation of type-1 and type-2 copper sites by addition of exogenous ligands to the *Pseudomonas aeruginosa* azurin His117Gly mutant." *J. Am. Chem. Soc.,* **1993**, *115*, 1121-1129.

(9) Di Bilio, A. J.; Chang, T. K.; Malmström, B. G.; Gray, H. B.; Karlsson, B. G.; Nordling, M.; Pascher, T.; Lundberg, L. G. "Electronic absorption spectra of M(II)(Met121X) azurins (M=Co, Ni, Cu; X=Leu, Gly, Asp, Glu): charge-transfer energies and reduction potentials." *Inorg. Chim. Acta,* **1992**, *198-200*, 145-148.

(10) Faham, S.; Mizoguchi, T. J.; Adman, E. T.; Gray, H. B.; Richards, J. H.; Rees, D. C. "Role of the active-site cysteine of *Pseudomonas aeruginosa* azurin. Crystal structure analysis of the CuII(Cys112Asp) protein." *J. Biol. Inorg. Chem.,* **1997**, *2*, 464-469.

(11) Karlsson, B. G.; Aasa, R.; Malmström, B. G.; Lundberg, L. G. "Rack-induced bonding in blue copper proteins: Spectroscopic properties and reduction potential of the azurin mutant Met-121 --> Leu." *FEBS Lett.,* **1989**, *253*, 99-102.

(12) Li, H.; Webb, S. P.; Ivanic, J.; Jensen, J. H. "Determinants of the Relative Reduction Potentials of Type-1 Copper Sites in Proteins." *J. Am. Chem. Soc.,* **2004**, *126*, 8010-8019.

(13) Pascher, T.; Karlsson, B. G.; Nordling, M.; Malmström, B. G.; Vänngård, T. "Reduction potentials and their pH dependence in site-directed-mutant forms of azurin from Pseudomonas aeruginosa." **1993**, *212*, 289-96.

(14) Crane, B. R.; Di Bilio, A. J.; Winkler, J. R.; Gray, H. B. "Electron Tunneling in Single Crystals of *Pseudomonas aeruginosa* Azurins." *J. Am. Chem. Soc.,* **2001**, *123*, 11623-11631.

(15) Farver, O.; Pecht, I. "Long range intramolecular electron transfer in azurins." *J. Am. Chem. Soc.,* **1992**, *114*, 5764-7.

(16) Farver, O.; Skov, L. K.; Pascher, T.; Karlsson, B. G.; Nordling, M.; Lundberg, L. G.; Vaenngaard, T.; Pecht, I. "Intramolecular electron transfer in single-site-mutated azurins." *Biochemistry,* **1993**, *32*, 7317-7322.

(17) Hough, M. A.; Strange, R. W.; Hasnain, S. S. "Conformational variability of the Cu site in one subunit of bovine CuZn superoxide dismutase: the importance of mobility in the Glu119-Leu142 loop region for catalytic function." *J. Mol. Biol.,* **2000**, *304*, 231-241.

(18) Solomon, E. I.; Randall, D. W.; Glaser, T. "Electronic structures of active sites in electron transfer metalloproteins: contributions to reactivity." *Coord. Chem. Rev.,* **2000**, *200-202*, 595-632.

(19) Berry, S. M.; Ralle, M.; Low, D. W.; Blackburn, N. J.; Lu, Y. "Probing the Role of Axial Methionine in the Blue Copper Center of Azurin with Unnatural Amino Acids." *J. Am. Chem. Soc.,* **2003**, *125*, 8760-8768.

(20) Garner, D. K.; Vaughan, M. D.; Hwang, H. J.; Savelieff, M. G.; Berry, S. M.; Honek, J. F.; Lu, Y. "Reduction Potential Tuning of the Blue Copper Center in *Pseudomonas aeruginosa* Azurin by the Axial Methionine as Probed by Unnatural Amino Acids." *J. Am. Chem. Soc.,* **2006**, *128*, 15608-15617.

(21) LaCroix, L. B.; Shadle, S. E.; Wang, Y.; Averill, B. A.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. "Electronic Structure of the Perturbed Blue Copper Site in Nitrite Reductase; Spectroscopic Properties, Bonding, and Implications for the Entatic/Rack State." *J. Am. Chem. Soc.,* **1996**, *118*, 7755-7768.

(22) Solomon, E. I. "Spectroscopic Methods in Bioinorganic Chemistry: Blue to Green to Red Copper Sites." *Inorg. Chem.,* **2006**, *45*, 8012-8025.

(23) Lu, Y. In *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology*; McCleverty, J. A., Meyer, J. J., Eds.; Elsevier: Oxford, UK, 2004; Vol. 8, p 91-122.

(24) Lu, Y.; LaCroix, L. B.; Lowery, M. D.; Solomon, E. I.; Bender, C. J.; Peisach, J.; Roe, J. A.; Gralla, E. B.; Valentine, J. S. "Construction of a blue copper site at the native zinc site of yeast copper-zinc superoxide dismutase." *J. Am. Chem. Soc.,* **1993**, *115*, 5907-5918.

(25) Arciero, D. M.; Pierce, B. S.; Hendrich, M. P.; Hooper, A. B. "Nitrosocyanin, a Red Cupredoxin-like Protein from *Nitrosomonas europaea*." *Biochemistry.,* **2002**, *41*, 1703-1709.

(26) Basumallick, L.; Sarangi, R.; DeBeer George, S.; Elmore, B.; Hooper, A. B.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. "Spectroscopic and Density Functional Studies of the Red Copper Site in Nitrosocyanin; Role of the Protein in Determining Active Site Geometric and Electronic Structure." *J. Am. Chem. Soc.,* **2005**, *127*, 3531-3544.

(27) Basumallick, L.; Szilagyi, R. K.; Zhao, Y.; Shapleigh, J. P.; Scholes, C. P.; Solomon , E. I. "Spectroscopic Studies of the Met182Thr Mutant of Nitrite Reductase: Role of the Axial Ligand in the Geometric and Electronic Structure of Blue and Green Copper Sites." *J. Am. Chem. Soc.,* **2003**, *125*, 14784-14792.

(28) Lieberman, R. L.; Arciero, D. M.; Hooper, A. B.; Rosenzweig, A. C. "Crystal Structure of a Novel Red Copper Protein from *Nitrosomonas europaea*." *Biochemistry,* **2001**, *40*, 5674-5681.

(29) Whittaker, M.; Bergmann, D.; Arciero, D.; Hooper, A. B. "Electron transfer during the oxidation of ammonia by the chemolithotrophic bacterium *Nitrosomonas europaea*." *Biochim. Biophys. Acta,* **2000**, *1459*, 346-355.

(30) Gray, H. B.; Malmström, B. G.; Williams, R. J. P. "Copper Coordination in Blue Proteins." *J. Inorg. Biochem.,* **2000**, *5*, 551-559.

(31) Guckert, J. A.; Lowery, M. D.; Solomon , E. I. "Electronic Structure of the Reduced Blue Copper Active Site: Contributions to Reduction Potentials and Geometry." *J. Am. Chem. Soc.,* **1995**, *117*, 2817-2844.

(32) Kofman, V.; Farver, O.; Pecht, I.; Goldfarb, D. "Two-Dimensional Pulsed EPR Spectroscopy of the Copper Protein Azurin." *J. Am. Chem. Soc.,* **1996**, *118*, 1201-1206.

(33) Murphy, L. M.; Strange, R. W.; Karlsson, B. G.; Lundberg, L. G.; Pascher, T.; Reinhammar, B.; Hasnain, S. S. "Structural characterization of azurin from *Pseudomonas aeruginosa* and some of its methionine-121 mutants." *Biochemistry,* **1993**, *32*, 1965-1975.

(34) Olsson, M. H. M.; Ryde, U. "The influence of axial ligands on the reduction potential of blue copper proteins." *J. Biol. Inorg. Chem.,* **1999**, *4*, 654-663.

(35) Solomon, E. I.; Hedman, B.; Hodgson, K. O.; Dey, A.; Szilagyi, R. K. "Ligand K-edge X-ray absorption spectroscopy: covalency of ligand-metal bonds." *Coord. Chem. Rev.,* **2005**, *249*, 97-129.

(36) Solomon, E. I.; Penfield, K. W.; Gewirth, A. A.; Lowery, M. D.; Shadle, S. E.; Guckert, J. A.; LaCroix, L. B. "Electronic structure of the oxidized and reduced blue copper sites: contributions to the electron transfer pathway, reduction potential, and geometry." *Inorg. Chim. Acta,* **1996**, *243*, 67-78.

(37) van Pouderoyen, G.; Andrew, C. R.; Loehr, T. M.; Sanders-Loehr, J.; Mazumdar, S.; Hill, H. A. O.; Canters, G. W. "Spectroscopic and Mechanistic Studies of Type-1 and Type-2 Copper Sites in *Pseudomonas aeruginosa* Azurin As Obtained by Addition of External Ligands to Mutant His46Gly." *Biochemistry,* **1996**, *35*, 1397-1407.

(38) Karlsson, B. G.; Tsai, L.-C.; Nar, H.; Sanders-Loehr, J.; Bonander, N.; Langer, V.; Sjolin, L. "X-ray Structure Determination and Characterization of the *Pseudomonas aeruginosa* Azurin Mutant Met121Glu." *Biochemistry,* **1997**, *36*, 4089-4095.

(39) Nar, H.; Messerschmidt, A.; Huber, R.; van de Kamp, M.; Canters, G. W. "Crystal structure analysis of oxidized *Pseudomonas aeruginosa* azurin at pH 5.5 and pH 9.0 : A pH-induced conformational transition involves a peptide bond flip." *J. Mol. Biol.,* **1991**, *221*, 765-772.

(40) Romero, A.; Hoitink, C. W. G.; Nar, H.; Huber, R.; Messerschmidt, A.; Canters, G. W. "X-ray Analysis and Spectroscopic Characterization of M121Q Azurin : A Copper Site Model for Stellacyanin." *J. Mol. Biol.,* **1993**, *229*, 1007-1021.

(41) Yamakura, F.; Sugio, S.; Hiraoka, B. Y.; Ohmori, D.; Yokota, T. "Pronounced Conversion of the Metal-Specific Activity of Superoxide Dismutase from *Porphyromonas gingivalis* by the Mutation of a Single Amino Acid (Gly155Thr) Located Apart from the Active Site." *Biochemistry,* **2003**, *42*, 10790-10799.

(42) Abriata, L. A.; Ledesma, G. N.; Pierattelli, R.; Vila, A. J. "Electronic Structure of the Ground and Excited States of the CuA Site by NMR Spectroscopy." *J. Am. Chem. Soc.,* **2009**, *131*, 1939-1946.

(43) Bertini, I.; Fernandez, C. O.; Karlsson, B. G.; Leckner, J.; Luchinat, C.; Malmstrom, B. G.; Nersissian, A. M.; Pierattelli, R.; Shipp, E.; Valentine, J. S.; Vila, A. J. "Structural Information through NMR Hyperfine Shifts in Blue Copper Proteins." *J. Am. Chem. Soc.,* **2000**, *122*, 3701-3707.

(44) Donaire, A.; Jimenez, B.; Fernandez, C. O.; Pierattelli, R.; Niizeki, T.; Moratal, J.-M.; Hall, J. F.; Kohzuma, T.; Hasnain, S. S.; Vila, A. J. "Metal-Ligand Interplay in Blue Copper Proteins Studied by 1H NMR Spectroscopy: Cu(II)-Pseudoazurin and Cu(II)-Rusticyanin." *J. Am. Chem. Soc.,* **2002**, *124*, 13698-13708.

(45) Kroes, S. J.; Hoitink, C. W. G.; Andrew, C. R.; Sanders-Loehr, J.; Messerschmid, A.; Hagen, W. R.; Canters, G. W. "The mutation Met121-->His creates a type1.5 copper site in *Alcaligene denitrificans* azurin." *Eur. J. Biochem.,* **1996**, *240*, 342-351.

(46) Messerschmidt, A.; Prade, L.; Kroes, S. J.; Sanders-Loehr, J.; Huber, R.; Canters, G. W. "Rack-Induced Metal Binding Vs. Flexibility - Met121his Azurin Crystal Structures At Different pH." *Proc. Natl. Acad. Sci. USA.,* **1998**, *95*, 3443-3448.

(47) Van Amsterdam, I. M. C.; Ubbink, M.; Van den Bosch, M.; Rotsaert, F.; Sanders-Loehr, J.; Canters, G. W. "A new type-2 copper cysteinate azurin: The involvement of an engineered exposed cysteine in copper binding through internal rearrangement." *J. Biol. Chem.,* **2002**, *277*, 44121-44130.

(48) van Gestel, M.; Boulanger, M. J.; Canters, G. W.; Huber, M.; Murphy, M. E. P.; Verbeet, M. P.; Groenen, E. J. J. "A Single-Crystal Electron Paramagnetic Resonacne Study at 95 GHz of the Type 1 Copper Site of the Green Nitrite Reductase of *Alcaligene faecalis*." *J. Phys. Chem,* **2001**, *105*, 2236-2243.

(49) Strange, R. W.; Murphy, L. M.; Karlsson, B. G.; Reinhammar, B.; Hasnain, S. S. "Effect of pH and Ligand Binding of the Structure of the Cu Site of the Met121Glu Mutant of Azurin from *Pseudomonas aeruginosa*." *Biochemistry,* **1996**, *35*, 16391-16398.

(50) Webb, M. A.; Kiser, C. N.; Richards, J. H.; Di Bilio, A. J.; Gray, H. B.; Winkler, J. R.; Loppnow, G. R. "Resonance Raman Spectroscopy of Met121Glu Azurin." *J. Phys. Chem.,* **2000**, *104*, 10915-10920.

(51) Berry, S. M.; Gieselman, M. D.; Nilges, M. J.; van der Donk, W. A.; Lu, Y. "An Engineered Azurin Variant Containing a Selenocysteine Copper Ligand." *J. Am. Chem. Soc.,* **2002**, *124*, 2084-2085.

(52) Chang, T. K.; Iverson, S. A.; Rodrigues, C. G.; Kiser, C. N.; Lew, A. Y. C.; Germanas, J. P.; Richards, J. H. "Gene Synthesis, Expression, and Mutagenesis Of the Blue Copper Proteins Azurin and Plastocyanin." *Proc. Natl. Acad. Sci. USA.,* **1991**, *88*, 1325-1329.

(53) Karlsson, B. G.; Nordling, M.; Pascher, T.; Tsai, L.-C.; Sjolin, L.; Lundberg, L. G. "Cassette mutagenesis of Met121 in azurin from *Pseudomonas aeruginosa*." *Protein Eng.,* **1991**, *4*, 343-349.

(54) Kolczak, U.; Dennison, C.; Messerschmid, A.; Canters, G. W. In *Handbook of Metalloproteins*; Messerschmid, A., Huber, R., Poulos, T., Wieghardt, K., Eds.; Wiley: Chichester, 2001; Vol. 2, p 1170-1194.

(55) den Blaauwen, T.; van de Kamp, M.; Canters, G. W. "Type I and II copper sites obtained by external addition of ligands to a His117Gly azurin mutant." *J. Am. Chem. Soc.,* **1991**, *113*, 5050-5052.

(56) Jeuken, L. J. C.; van Vliet, P.; Verbeet, M. P.; Camba, R.; McEvoy, J. P.; Armstrong, F. A.; Canters, G. W. "Role of the Surface-Exposed and Copper-Coordinating Histidine in Blue Copper Proteins: The Electron-Transfer and Redox-Coupled Ligand Binding Properties of His117Gly Azurin." *J. Am. Chem. Soc.,* **2000**, *122*, 12186-12194.

(57) Lu, Y.; Roe, J. A.; Bender, C. J.; Peisach, J.; Banci, L.; Bertini, I.; Gralla, E. B.; Valentine, J. S. "New type 2 copper-cysteinate proteins. Copper site histidine-to-cysteine mutants of yeast copper-zinc superoxide dismutase." *Inorg. Chem.,* **1996**, *35*, 1692-700.

(58) Mizoguchi, T. J.; Di Bilio, A. J.; Gray, H. B.; Richards, J. H. "Blue to Type 2 Binding. Copper(II) and Cobalt(II) Derivatives Of a Cys112Asp Mutant Of *Pseudomonas aeruginosa* Azurin." *J. Am. Chem. Soc.,* **1992**, *114*, 10076-10078.

(59) Ralle, M.; Berry, S. M.; Nilges, M. J.; Gieselman, M. D.; van der Donk, W. A.; Lu, Y.; Blackburn, N. J. "The Selenocysteine-Substituted Blue Copper Center: Spectroscopic Investigations of Cys112SeCys *Pseudomonas aeruginosa* Azurin." *J. Am. Chem. Soc.,* **2004**, *126*, 7244-7256.

(60) van Gestel, M.; Canters, G. W.; Krupka, H.; Messerschmid, A.; de Waal, E. C.; Warmerdam, G. C. M.; Groenen, E. J. J. "Axial Ligation in Blue-Copper Proteins. A W-Band Electron Spin Echo Detected Electron Paramagnetic Resonance Study of the Azurin Mutant M121H." *J. Am. Chem. Soc.,* **2000**, *122*, 2322-2328.

(61) Hay, M.; Richards, J. H.; Lu, Y. "Construction and characterization of an azurin analog for the purple copper site in cytochrome *c* oxidase." *Proc. Natl. Acad. Sci. USA.,* **1996**, *93*, 461-464.

(62) Randall, D. W.; Gamelin, D. R.; LaCroix, L. B.; Solomon , E. I. "Electronic structure contributions to electron transfer in blue Cu and CuA." *J. Biol. Inorg. Chem.,* **2000**, *5*, 16-29.

(63) Sieracki, N. A.; Hwang, H. J.; Lee, M. K.; Garner, D. K.; Lu, Y. "A temperature independent pH (TIP) buffer for biomedical biophysical applications at low temperatures." *Chem. Comm.,* **2008**, 823-825.

(64) David, R.; Richter, M. P. O.; Beck-Sickinger, A. G. "Expressed protein ligation. Method and applications." *Eur. J. Biochem.,* **2004**, *271*, 663-677.

(65) Muir, T. W.; Sondhi, D.; Cole, P. A. "Expressed protein ligation: A general method for protein engineering." *Proc. Natl. Acad. Sci. USA.,* **1998**, *95*, 6705-6710.

(66) Johnson, E. C. B.; Kent, S. B. H. "Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction." *J. Am. Chem. Soc.,* **2006**, *128*, 6640-6646.

(67) Cheung, K. C.; Strange, R. W.; Hasnain, S. S. "3D EXAFS refinement of the Cu site of azurin sheds light on the nature of structural change at the metal centre in an oxidation-reduction process: an integrated approach combining EXAFS and crystallography." *Acta. Crystallogr. D Biol. Crystallogr.,* **2000**, *56*, 697-704.

(68) Andruzzi, L.; Nakano, M.; Nilges, M. J.; Blackburn, N. J. "Spectroscopic studies of metal binding and metal selectivity in *Bacillus subtilis* BSco, a homologue of the yeast mitochondrial protein Sco1p." *J. Am. Chem. Soc.,* **2005**, *127*, 16548-16558.

CHAPTER 5

PROBING BACKBONE CARBONYL INTERACTIONS OF CU_A AZURIN.

5.1 INTRODUCTION

Engineering and exploring the function of metalloproteins requires intimate knowledge of the metal binding residues. Without establishing specific function of residues, further optimization or protein design becomes difficult. Traditionally, the function of metal binding residues is investigated using site-directed mutagenesis (SDM) with proteinogenic amino acids. As discussed in previous chapters, more recently, the incorporation of nonproteinogenic amino acid analogues has been favored over SDM in the study of specific residue function because the dramatic geometric and electronic changes associated with SDM are constrained using nonproteinogenic amino acids. Expressed Protein Ligation (EPL) has previous been used to incorporate nonproteinogenic amino acids into metalloprotein scaffolds to clarify residue function and further aid attempts at protein engineering.[1-6] The Cu_A azurin scaffold however has not been subjected to analysis using EPL.

As discussed in chapter 1, the Cu_A site geometry is different than its electron transfer cousins, the type 1 copper proteins.[7] Its geometry is more similar to dinuclear non-heme iron proteins than to other Cu electron transfer sites (See Figure 1.8). The two Cu ions sit in a rigid diamond core consisting of μ2 bridging thiolates from two cysteine residues (Figure 5.1).[8-15] The Cu ions are coordinated by the δ nitrogens from two His imidazole rings and two additional axial ligands.[8-11,13-15] Typically these axial ligands are Met thioethers and backbone carbonyl groups of nearby residues.[13-15]

As described previously in chapter 1, the study of the Cu_A site in its native scaffold is challenging since the native proteins contain multiple other metal-binding sites giving rise to cross signals that interfere with data collection and make interpretation difficult. To alleviate these troubles, the Cu_A azurin scaffold was created to easily express a stable spectroscopic mimic of native Cu_A sites. The engineered site mimics both the geometry and spectroscopic signatures of natural Cu_A sites.[10,13,14,23,24] The protein has LMCT bands at 480 nm and 530 nm resulting from the S(Cys) → Cu(II) charge transfer transition.[13,14,19,20] An additional *d-d* transition between the Cu ions results in a third LMCT band at 775 nm. The site also has a small $A_{||}$ value in the EPR spectrum from its delocalized electrons, resulting in a 7-line hyperfine splitting (See Figure 1.9).[13,14,19,20]

Thus far substitution of residues in metalloproteins with nonproteinogenic amino acids has probed the function of ligands that use a side chain to bind Cu.[1-3,6] While researchers are intensely interested in substituting Cu ligands in Cu_A azurin with nonproteinogenic amino acid side-chain analogues, efforts are lacking with respect to elucidating the function of the unique backbone carbonyl of Glu114 that is a key part of the Cu coordination in the site. As seen in Figure 5.1, one copper in the active site is bound by the backbone carbonyl of Glu114. Replacing the Glu114 amide moiety with isosteres could help to establish a new set of data that better explains backbone carbonyl contribution in Cu coordination. While the synthesis of backbone analogues can be more straightforward than the synthesis of nonproteinogenic amino acids with altered side chain functionality, peptide synthesis using backbone analogues is more difficult with respect to standard solid-phase peptide synthesis (SPPS) strategies because forming the peptide bond is crucial to peptide elongation. Three isosteres of peptide bonds have been

incorporated into peptide chains: 1) amide-to-ester, 2) amide-to-thioamide, and 3) amide-to-amine isosteres. Incorporating these three isosteres into Cu_A azurin, may provide new insights into the Cu-to-carbonyl interaction.

5.2 EPL OF WT Cu_A AZURIN

As a control prior to synthesizing any peptides containing nonproteinogenic amino acid residues, the WT Cu_A azurin sequence (N_2H-CSELCGINHALMKGTLTLK-COOH) consisting of the final 19 residues in the Cu_A azurin protein was synthesized following standard SPPS strategies using Fmoc-Cys(Trt)-OH as the Cys116 derivative (Peptide ***5.1,*** Figure 5.1). The peptide was cleaved from standard Wang resin using a cocktail of 2% H_2O/TIPS/EDT in neat TFA at RT for 2 h. Following synthesis and purification of the peptide ***5.1***, EPL reactions were run. Samples were taken at 24 h, 48 h, and 66 h and checked for ligation progress. However, even in the presence of excess TCEP, the ligations failed to proceed resulting only in hydrolyzed azurin-thioester and free peptide.

Inefficient ligation was attributed to potential disulfide bond formation. Cys112 and Cys116 are three residues apart, on the same face of the peptide, and are likely to form an intramolecular disulfide bond. This intramolecular disulfide bond impedes participation of the Cys112 thiol in the transthioesterification of the peptide with the azurin-thioester. Thus, the efficiency of the ligation is poor resulting in extensive time for hydrolysis of the azurin-thioester and no ligated product. To prevent inefficient ligation an alternative peptide was synthesized replacing Cys116 with Fmoc-Cys(StBu)-OH (N_2H-CSELC$_{(StBu)}$GINHALMKGTLTLK-COOH) (Peptide ***5.2*** Figure 5.1). The StBu

protected Cys116 prevents intramolecular disulfide formation and increases ligation efficiency. However, use of TCEP is limited because it breaks the Cys(StBu) disulfide. Thus any intermolecular disulfides between Cys112 residues cannot be reduced except through disulfide exchange and reduction with the thiol catalyst, sodium 2-mercaptoethanesulfonate (MESNa).

Peptide ***5.2*** was synthesized using standard Fmoc-based SPPS strategies and cleaved from standard Wang resin in the conditions already described. EPL reactions were carried out with the Cys116StBu protected peptide in the presence of MESNa and in the absence of TCEP. Samples taken after 24 h, 48 h and 66 h all contained correctly ligated protein with fully ligated protein yields increasing with time. The EPL eluent was purified by cation exchange chromatography, resulting in two separate protein isoforms; one with a free Cys116 thiolate (likely from reduction by MESNa) and another with the Cys116StBu disulfide intact as determined my MALDI-MS.

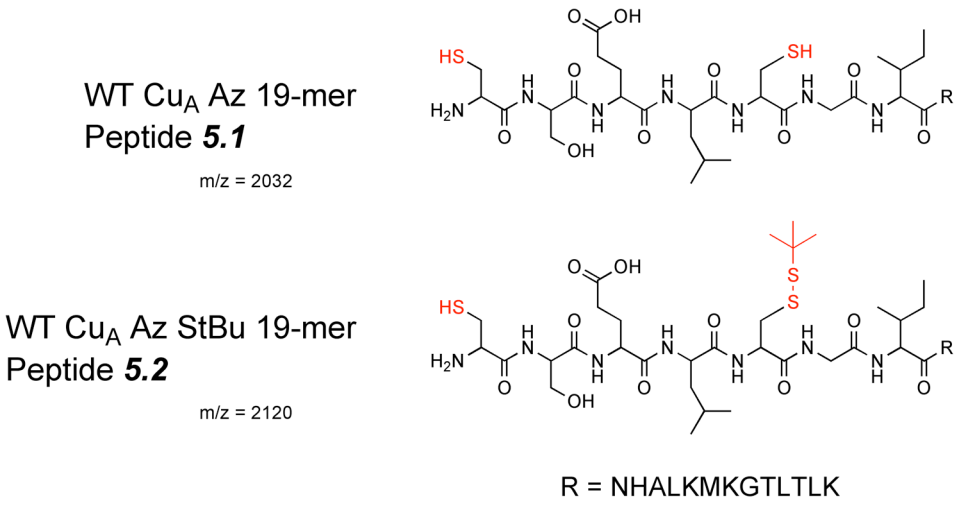


Figure 5.1: 19-mer peptides synthesized for EPL reactions for WT Cu_A azurin. The free thiolate on Cys116 results in intramolecular disulfide bond formation and a lack of efficient transthioesterification in the EPL reaction.

The first fractions from the purification contained Full Length (FL) WT Cu_A Az correctly folded with both thiolates of Cys112 and Cys116 free. The combined fractions were titrated without any need for refolding. The protein correctly bound Cu(II) but at a R/Z ratio of 7, nearly double that of recombinant WT Cu_A Az (R/Z = 3), suggesting that the protein is not entirely correctly folded or that a portion of the protein is damaged and unable to bind Cu (Figure 5.2 A, D).[13] The second set of fractions contained FL WT Cu_A Az that was poorly folded as determined by the A_{291} peak.[28] The sharp shoulder on the A_{280} band is indicative of proper folding of azurin by the burying of Trp48 into the protein.[28] When the band is sharp, the protein is deemed well folded. When the band is broad, it indicates improper folding and the necessity to refold the apo protein.[28] Because the A_{291} peak for the second fraction set is broad, it is indicative of improper folding and may explain why the apo protein was unable to bind Cu since Cys116 remained protected by the StBu disulfide and was therefore unable to participate in Cu binding (Figure 5.2 B).[28]

To remove the remaining StBu protecting group bound to Cys116, the sample was refolded in the presence of dithiothreitol (DTT). The DTT treatment was able to remove all remaining StBu groups resulting in better folded protein that was able to bind Cu with a high R/Z ratio (R/Z = 3) (Figure 5.2 C, E). The R/Z value of the refolded protein was equivalent to correctly folded recombinant WT Cu_A Az and the shoulder at A_{291} was now present as a sharp peak (Figure 5.2 C).[13,14] This increase in Cu binding suggests that the protein is better folded than the first fraction set that contained FL WT Cu_A Az without the StBu group bound to Cys116. Thus, while it appeared not necessary to refold the first fraction set since examination of the A_{291} peak indicative of correct folding was sharp and

the protein correctly bound Cu(II), to increase Cu binding and thus the yield of holo-protein, in further studies the entire eluent after the EPL reactions was refolded prior to cation exchange purification. Refolding ensured that not only the Cys116 thiolate was free, but also that more protein was correctly folded before titrating with Cu(II).

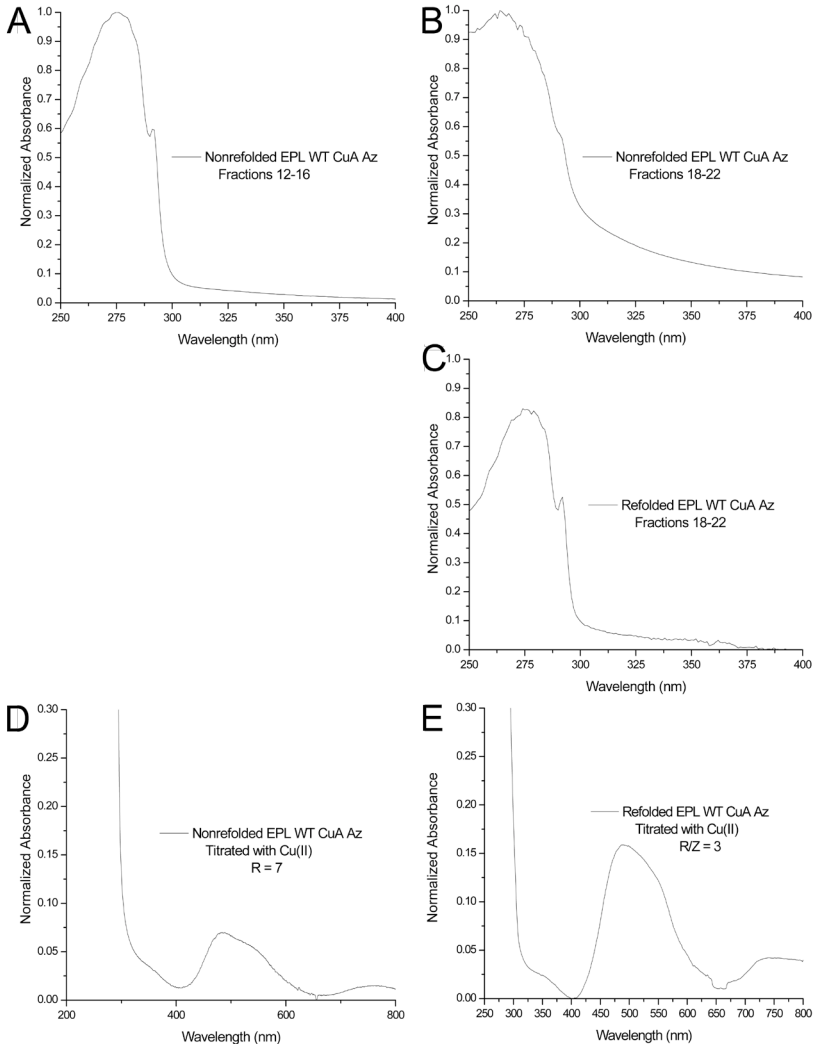


Figure 5.2: FL EPL of WT Cu_A azurin. A) First fraction set without any refolding steps and before titration. The fractions contain protein with free thiolates for Cys112 and Cys116. The peak at A_{291} is indicative of well-folded protein. B) Second fraction set without refolding steps and before titration. The fractions contain protein with the Cys116StBu intact. The fraction is poorly folded as seen by the broad A_{291} peak. C) Second fraction set after refolding and before titration. The protein now has free thiolates for Cys112 and Cys116 and is very well folded by a well-defined A_{291} peak. D) First fraction set without any refolding steps fully titrated with Cu at an R/Z value of 7 (WT R/Z = 3). E) Second fraction set after refolding and fully titrated with Cu at the WT R/Z ratio (R/Z = 3).

After refolding the semi-synthesized WT Cu_A Az, the resulting apo-protein was titrated with $Cu(II)SO_4$ to verify its spectroscopic properties matched recombinant WT Cu_A Az. All titrations of EPL product were completed at pH 5.1 in NH_4OAc buffer. Titration of the EPL produced WT Cu_A Az protein yielded a purple colored solution that deepened in color with increasing additions of Cu(II). Once fully titrated (at saturation of A_{480}), the buffer was exchanged to fresh 50 mM NH_4OAc and the UV-Vis spectra reexamined to verify the LMCT bands remained unchanged. As shown in Figure 5.3, the LMCT bands of EPL generated WT Cu_A Az at 480 nm, 530 nm, and 750 nm matched the LMCT bands of recombinant WT Cu_A Az (480 nm, 530 nm, 750 nm). X-band EPR also verified correct formation of the Cu_A site in EPL WT Cu_A Az.[13,14] Consequently, EPL generated WT Cu_A Az is spectroscopically indistinguishable from recombinant WT Cu_A Az.

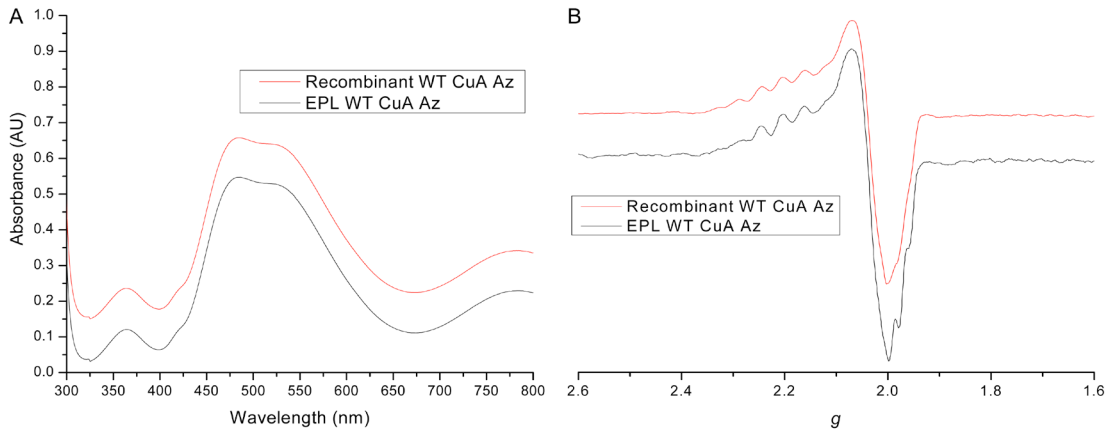


Figure 5.3: Spectroscopic comparison of EPL WT Cu_A azurin with recombinantly expressed WT Cu_A azurin. Both proteins were titrated with the same equiv. of Cu(II) and the spectra were normalized by concentration. A) UV-Vis spectroscopic comparison of EPL and recombinant WT Cu_A Az shows very similar LMCT bands. B) X-band EPR analysis of EPL and recombinant WT Cu_A Az shows 7-line hyperfine splitting indicative of Cu_A sites with identical *g* values.

5.3 Synthesis of Backbone Analogue Peptides

Two analogues were chosen for investigating the Cu-carbonyl interaction: 1) amide-to-ester replacement and 2) amide-to-thioamide replacement (Figure 5.4). These two analogues are very similar in size and thus will likely minimally perturb the geometry of the site. The analogues also permit precise study of the Cu-carbonyl electronic interaction. Additionally, the amide-to-ester and amide-to-thioamide replacements alter the basicity of the carbonyl potentially reducing or increasing the electron density of the carbonyl and altering its interaction with the Cu.

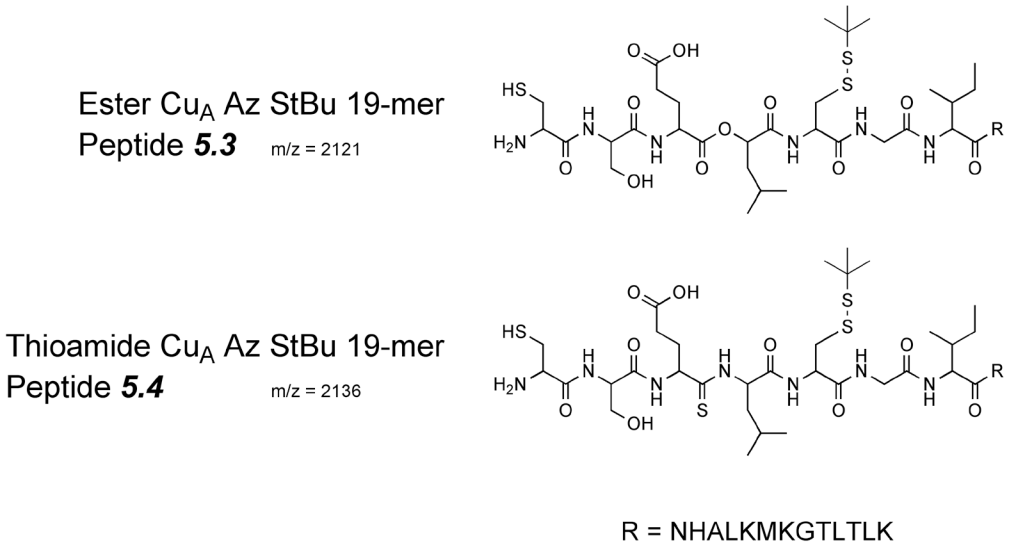


Figure 5.4: Peptides for investigation of the Cu-carbonyl interaction in Cu_A azurin. All peptides contain the Cys116StBu protection to prevent intramolecular disulfide formation.

Design of nonproteinogenic amino acid analogues for metal ligand substitution in metalloproteins typically focuses on altering side chain functionality retaining both the amine and carboxylate functionalities necessary for amide synthesis in standard SPPS chemistry.[1-3,6] However, since it is the carboxylate on the backbone analogues that is modified, alternative Fmoc-based SPPS strategies are necessary to form linkages in

peptides containing backbone carbonyl analogues. Installing an ester mimics installing an amide so very little deviation from standard SPPS strategies is required; only different coupling agents are necessary.[29,30] Similarly, installing the thioamide does not deviate from standard SPPS strategies; the thioamide moiety is built with the activating agent in place.[31] However, installing a reduced isostere moiety in replacement of an amide requires chemistry not typically found in standard SPPS schemes.[32]

5.4 Installing an Ester Into Cu_A Azurin

Installing ester linkages into peptides creates chains termed depsipeptides. Any number of ester linkages into a peptide sequence can be made. Linear synthesis of depsipeptides is well established in the literature and the current interest in depsipeptides has prompted exploration of methodologies for automated syntheses.[29,30,33-36] Initially depsipeptides were synthesized as peptidomimetics of antimicrobials that reportedly had increased stability and protection against degradation by proteases.[36] Because of the changes in H-bonding properties of ester linkages when compared to amides in peptides, depsipeptides have also been used to study α-helical H-bonding networks and have aided in the study of ion transporters.[4,5,34,35]

Synthesis of depsipeptides involves extensive optimization to reduce racemization typically found in protocols for ester synthesis.[4,5,36] Additionally, esters are more susceptible to base hydrolysis than their amide counterparts. Consequently, depsipeptides are known to fragment in the presence of strong bases.[4,36] Moreover, forming esters is a slower process than forming amides and requires increased reaction times to ensure complete coupling.[4,5,36] Racemization, fragmentation, and long reaction times associated

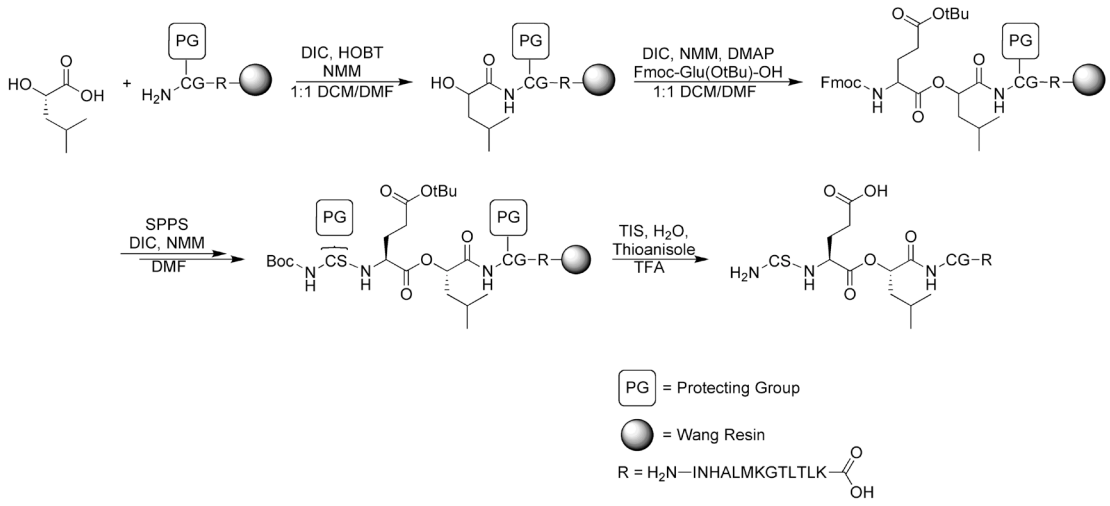
with installing esters resulted in development of alternative routes to synthesize depsipeptides using standard Fmoc-based SPPS. Many available schemes limit exposure of synthesized depsipeptides to the strong bases normally used for Fmoc-deprotection (piperidine) in Fmoc-based SPPS by using other bases such as 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU).[4,5,29,30,37,38] While decreasing potential base hydrolysis of installed esters, using different bases does not eliminate racemization and the long reaction times associated with ester synthesis. Less than 1% racemization after ester coupling is present when esters in depsipeptide sequences are installed using diisopropylcarbodiimide (DIC)/4-dimethylaminopyridine (DMAP) chemistry.[4,5,36] Additionally, DIC/DMAP chemistry was shown to be extremely efficient at installing esters in resin bound peptides, reducing coupling times to approximately 1-2 h.

However, installation of ester linkages in peptides sequences while simultaneously reducing base hydrolysis, racemization and maintaining short coupling times is extremely dependent on the peptide sequence. Thus when installing ester moieties in peptides, every sequence must be optimized to limit byproducts.[4,36] Sequences that contain regions prone to α-helix formation severely limit the efficiency of installing esters in the absence of chaotropic agents such as trifluoroethanol (TFE).[30,36] Additionally, the presence of acidic residues such as Glu and Asp can result in cyclization and hydrolysis of the ester.[4,36]

5.5 SPPS OF PEPTIDE *5.3* BY FMOC CHEMISTRY

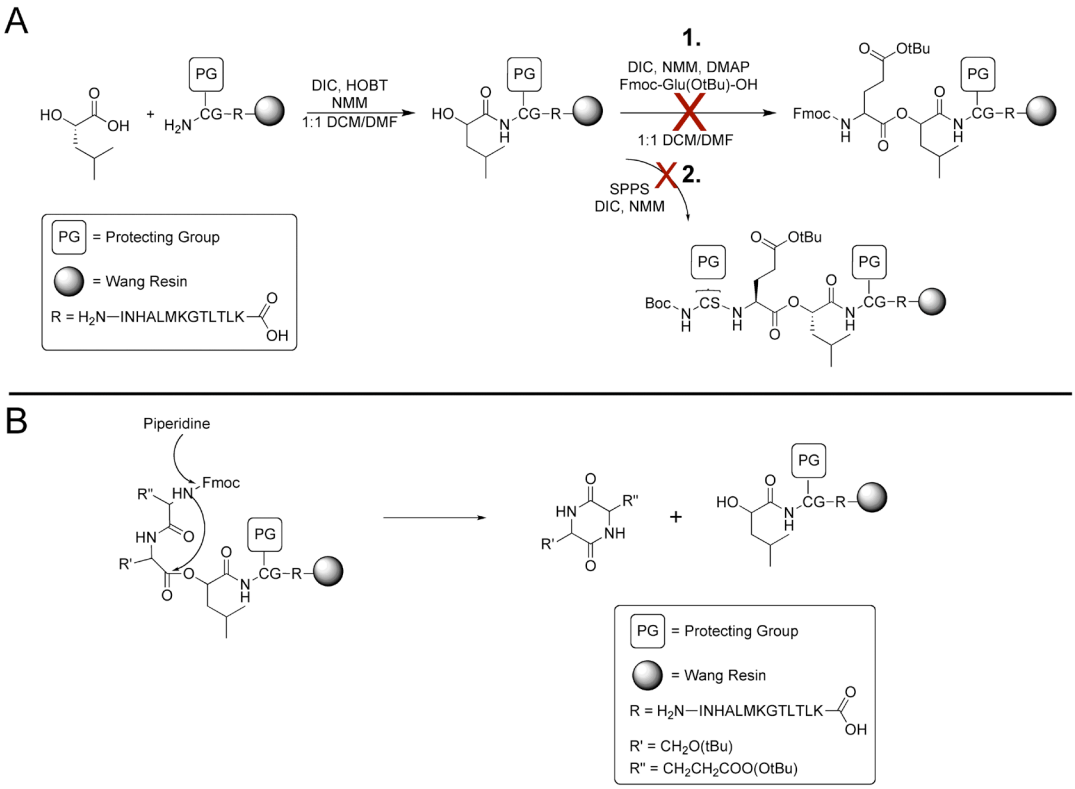
Synthesis of a 19-mer peptide (H_2N-CSE[ΨCOO]$LC_{(StBu)}$GINHALMKGTLTLK-OH) containing an ester between residues Glu114 and Leu115 was required for ester

installation. Initial syntheses using Cys116 as the free thiolate instead of the StBu disulfide resulted in inefficient ligation similar to the observations in WT ligations. Thus SPPS of the peptide ***5.3*** was completed using Cys116 protected with the StBu disulfide to prevent intramolecular disulfide formation. Installing the ester required a different strategy from the SPPS of peptide ***5.2***. Forming an ester bond is slower than forming an amide bond since the hydroxyl group is less nucleophilic than an amine.[4,39] Because the reaction is slower, coupling times for forming the ester were longer than those for amide bond couplings. To minimize racemization, no base was used during the coupling to install the ester and DIC/DMAP chemistry was used to speed the reaction (Scheme 5.1). The coupling was monitored using a hydroxyl specific test similar to the Kaiser Test (see experimental).[40] The two residues following Glu114 (Ser113 and Cys112) were coupled using standard HOBt/DIC chemistry.



Scheme 5.1: Fmoc-SPPS of the peptide ***5.3***. Standard HOBt/DIC coupling with Leucic acid and a pre-formed 15-mer generates the hydoxyl peptide for esterification. The ester is installed by DIC/DMAP chemistry with Fmoc-Glu(OtBu)-OH. Following Fmoc-deprotection, the final two amino acids are added using HOBt/DIC chemistry and cleaved in neat TFA with scavengers.

Initial peptide synthesis attempts used piperidine for Fmoc-deprotection of the residues following the ester at Glu114 (Ser113 and Cys112) and exclusively resulted in truncated peptide containing residues Leu115-Lys130. Two potential routes may have caused poor product yield. First, the esterification itself may be incomplete, leading to peptide truncated at Leu115 and the ensuing couplings of Ser113 and Cys112 would fail using HOBt/DIC chemistry (Scheme 5.2A).[4] A second potential problem is diketopiperazine (DKP) formation upon deprotection of Fmoc-Ser(OtBu) using piperidine. The cyclized DKP results in breakage of the ester linkage (Scheme 5.2B).[4]



Scheme 5.2: Potential explanations for failure of synthesis of FL peptide ***5.3***. A) Synthesis up to the installation of leucic acid is standard. Incomplete installation of the ester results in failure to complete route 1 and thus to install Glu114. This failure means that the hydroxyl of Leu115 remains and is therefore unable to participate in further SPPS to install Ser113 and Cys112 and FL peptide ***5.3*** is therefore unattainable. B) Diketopiperazine (DKP) formation occurs when the amine of Ser113 attacks the ester after Fmoc-deblocking with piperidine. Glu114 and Ser113 are cyclized into the piperazine ring, releasing the free hydoxyl peptide.

The first hypothesis that the esterification itself was inefficient was ruled out by testing with a hydroxyl specific colorimetric test similar to the Kaiser test.[40] The test is able to determine the presence of free hydroxyls at extremely low percentages (~1%).[36,40] After triple coupling of Glu114 no free hydroxyls were present as determined by the test. Furthermore, trial cleavage of the peptide determined that Glu114 was present with the Fmoc protecting group intact, also ruling out breakage of the ester by ionization. The hypothesis that DKP formation resulted in cleavage of the ester was tested by using a weaker base, 1% DBU at reduced times for Fmoc-deprotection of Glu114 and Ser113.[37] To further limit exposure to base, Cys112 was coupled as the Boc-protected amine that could be removed during cleavage of the peptide. Reduced exposure to base afforded increased peptide yields but retained significant hydrolyzed peptide byproduct (Figure 5.5). Therefore an alternative strategy, Boc-SPPS, was used.

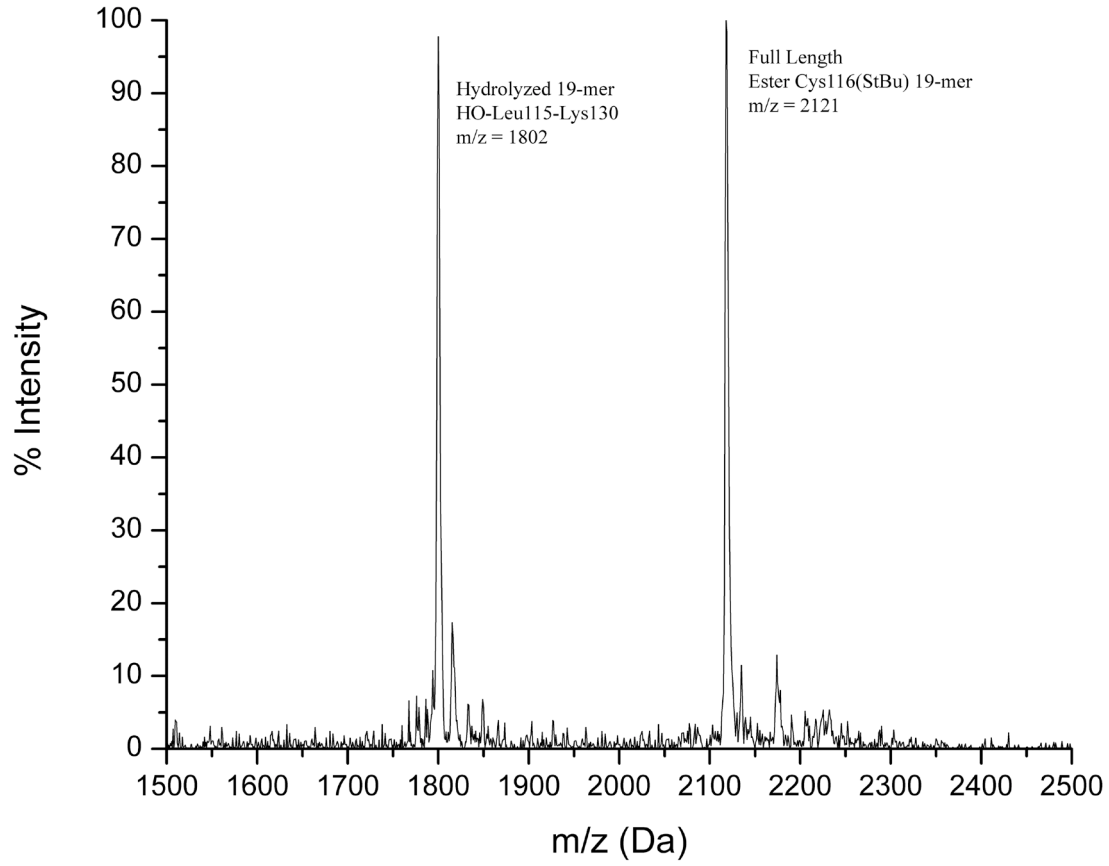
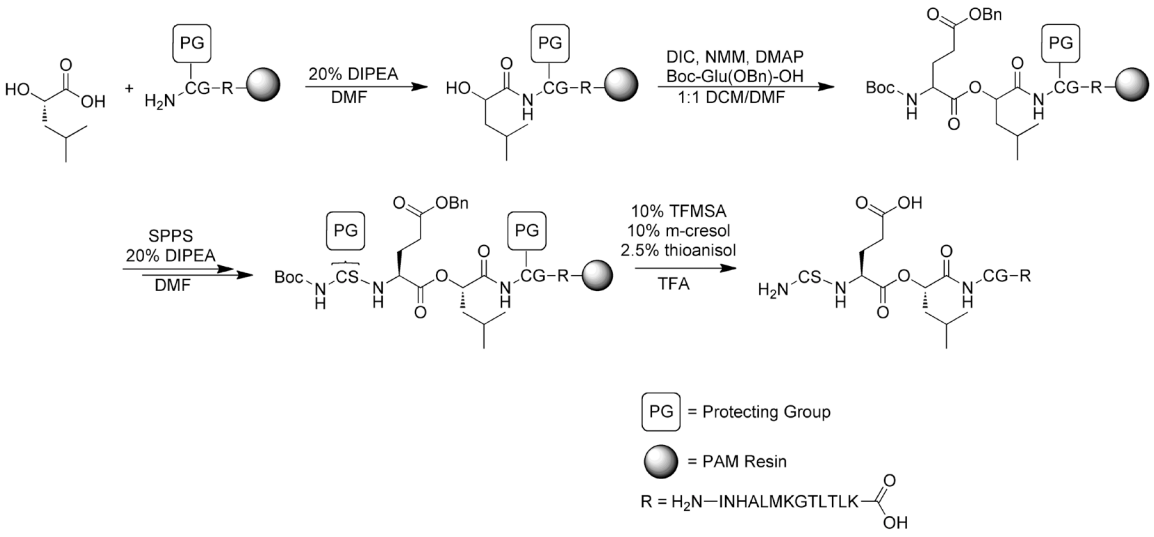


Figure 5.5: MALDI-MS analysis of the Peptide ***5.3*** synthesized *via* Fmoc-SPPS. DBU was used to deprotect residues after installation of the ester. Even in the presence of weaker base, DKP formation results, causing ester hydrolysis and resulting in a large proportion of truncated peptide with a mass of 1802 Da.

5.6 SPPS OF PEPTIDE *5.3* BY BOC CHEMISTRY

As an alternative to Fmoc-based SPPS, peptide ***5.3*** was synthesized using Boc-SPPS chemistry. Boc chemistry eliminates exposure to base, eliminating DKP formation and base hydrolysis. Peptide synthesis using Boc chemistry however is a much more harsh route to synthesize peptides because of the strong acids necessary during synthesis and for peptide cleavage from the resin. Though alternative routes for Boc removal exist other than using TFA, many of these strategies employ expensive reagents or require long deprotection times.[41-43] Resins exist that do not require HF for cleavage and instead use

superacids such as trimethylsilyl trifluoromethanesulfonate (TMSOTf) or trifluoromethanesulfonic acid (TFMSA) for cleavage.



Scheme 5.3: Boc-SPPS of the peptide ***5.3***. Standard coupling of leucic acid with 20% DIPEA prepares the hydroxyl peptide for esterification. The ester is installed by DIC/DMAP chemistry with Boc-Glu(OBn)-OH. Following Boc-deprotection, the final two amino acids are added using HOBt/DIC chemistry and the peptide was cleaved from the resin in neat TFA with TFMSA with scavengers. TMSOTf was shown to be insufficient to fully cleave the peptide.

Synthesis of peptide ***5.3*** using Boc chemistry was completed on H-Lys(2-Cl-Z)-PAM resin using the Boc amino acid derivatives listed in the experimental section (Scheme 5.3). Trial cleavages of a shorter peptide (H_2N-ALMKGTLTLK-OH) using TMSOTf or TFMSA were completed before the FL peptide was synthesized to verify that cleavage procedures were sufficient to fully cleave the peptide. Samples taken at 2 h and 4 h showed no peptide product from the TMSOTf mixture and cleavage was deemed insufficient. Cleavage of the shorter peptide with TFMSA showed fully cleaved peptide of the correct mass. Following trial cleavages, the synthesized FL peptide was completed

using Boc-Cys(StBu)-OH at position 116 to prevent intramolecular disulfide formation during future EPL ligations.

The ester was installed using DIC/DMAP chemistry identically to the methods used in Fmoc-based SPPS and the FL peptide was cleaved from the resin in a 10% TFMSA cocktail in TFA for 2-4 h. Precipitation of the FL peptide was difficult as extraneous organics in the cleavage cocktail prevented peptide trituration, instead yielding a golden oil. Therefore, prior to trituration, the cleavage cocktail was extracted three times with dH_2O and the aqueous layer triturated with cold Et_2O. This procedure was sufficient to obtain the peptide as a white powder. However, yields after lyophilization were low. Instead, the cleavage cocktail was taken up in THF/TFA and extracted with dH_2O. The aqueous layer was then extracted with Et_2O, resulting in peptide yields after lyophilization of the aqueous layer that corresponded to the expected yield of crude peptide (150-250 mg crude material for a 0.1 mmol synthesis).

MALDI-MS characterization of the peptide synthesized *via* Boc chemistry showed less than 5% hydrolyzed product and the correct m/z of the peptide ***5.3*** (Figure 5.6). Though purified yields *via* Boc-SPPS (50-65 mg peptide) were lower than typical purified yields from Fmoc-SPPS (50-100 mg peptide), the lower yields are likely a result of cleavage of the peptide from the resin throughout the synthesis, a known problem using TFA to remove N-terminal Boc protection during Boc-SPPS.[41,43] Additionally, purified peptide from Fmoc syntheses contained significant amounts of other peptide impurities, including the hydrolyzed byproduct, likely resulting in inflated yields when compared to the higher purity peptide obtained through Boc-SPPS. Thus, the yield of FL correct mass peptide from Boc chemistry was superior to the same scheme using Fmoc

chemistry. However, though yields are higher, peptides synthesized *via* Boc chemistry in our lab cannot be automated and thus all synthesis must be done manually.

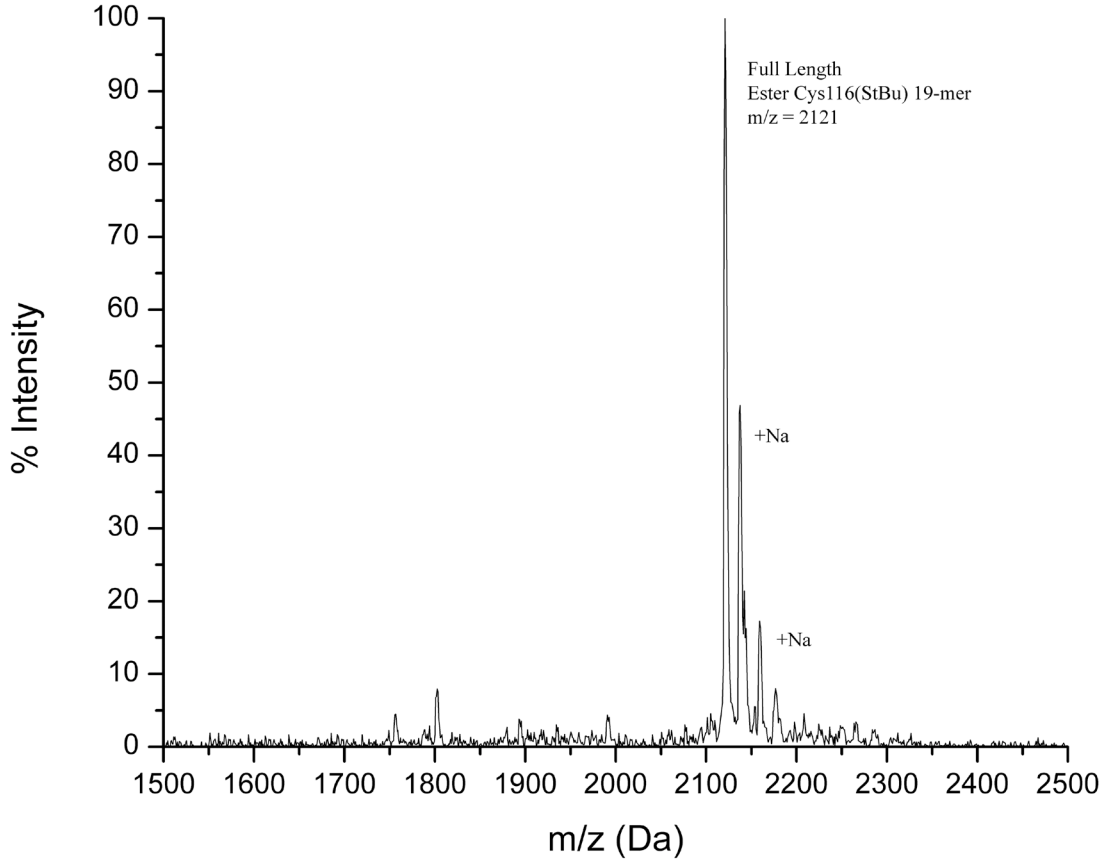


Figure 5.6: MALDI-MS analysis of peptide ***5.3*** synthesized via Boc-SPPS. Very little hydrolysis was noted and the peptide was clean prior to purification.

5.7 SEMI-SYNTHESIS OF ESTER-CU_A AZ BY EPL

The FL peptide ***5.3*** was purified by RP-HPLC and used in EPL reactions following the same methodology described for WT Cu_A Az reactions. Noteworthy however is the sensitivity of the ester linkage (both the peptide and the subsequent FL protein) in buffers at pH > 7.5. No ligated products were observed using buffers at pH > 7.5 during the course of the EPL reaction. MALDI-TOF MS analysis of EPL elutions in buffer at pH > 7.5 showed peptide that was fully hydrolyzed between Glu114 and Leu115

and was therefore unable to participate in transthioesterification with the protein-thioester. To increase the rates of transthioesterification and reduce hydrolysis, buffers at pH 7.2 were used for all ester Cu_A EPL reactions.

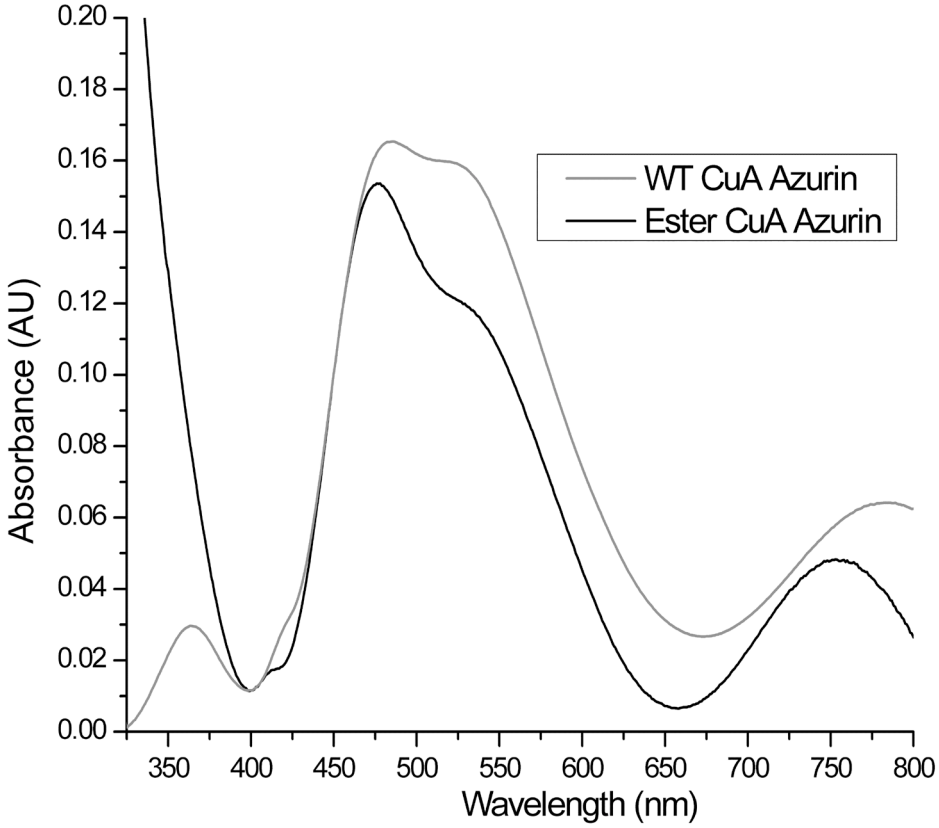


Figure 5.7: UV-Vis overlay of WT EPL Cu_A azurin and Ester Cu_A azurin. The LMCT bands are red shifted for the ester protein (475 nm, 530 nm, and 750 nm). The WT LMCT bands are at 480 nm, 530 nm and 775 nm.

Full length apo protein from EPL reactions was refolded immediately following ligation and titrated to 1 equiv. $Cu(II)SO_4$ where the protein solution turned purple in color. Holo-ester Cu_A azurin (hereafter ester Cu_A) showed three LMCT bands that corresponded very closely to, but were red-shifted from the ligand-to-metal charge transfer (LMCT) bands found in WT Cu_A azurin (hereafter WT) suggesting a stronger

carbonyl-Cu interaction (Figure 5.7). Ester Cu_A LMCT bands were noted at 475 nm, 530 nm, and 750 nm whereas the WT Cu_A LMCT bands are at 480 nm, 530 nm, and 775 nm. To verify that the ester Cu_A protein contained the ester moiety, hydrolysis of the protein by exposure to 1 *M* NaOH was performed. Because the ester is susceptible to base hydrolysis, exposure of the protein to 1 *M* NaOH will preferentially cleave the ester while the WT protein will remain intact. SDS-PAGE gel analysis was performed to observe the expected MW shifts of the cleaved ester protein. WT Cu_A Az (protein with a native amide bond) was not susceptible to hydrolysis using 1 *M* NaOH for 20 min at RT (Figure 5.8, lane 2). However, exposure of ester Cu_A protein to 1 *M* NaOH at 20 min and 60 min showed an incomplete hydrolysis product (lanes 4 and 5) that was not found in ester protein unexposed to 1 *M* NaOH (lane 3), verifying correct incorporation of the ester moiety.

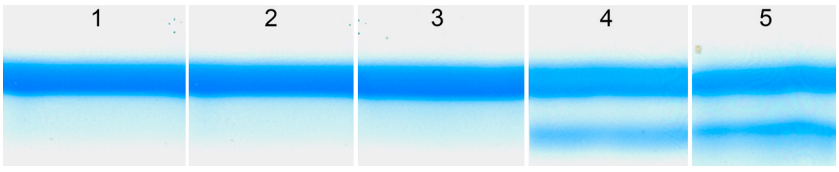


Figure 5.8: SDS-PAGE gel analysis of ester incorporation into Cu_A azurin. See text for details. Lane 1 = FL WT EPL Cu_A Az, Lane 2 = FL WT EPL Cu_A Az exposed to 1 *M* NaOH, 20 min., Lane 3 = FL Ester Cu_A Az, Lane 4 = FL Ester Cu_A Az exposed to 1 *M* NaOH, 20 min., Lane 5 = FL Ester Cu_A Az exposed to 1 *M* NaOH, 60 min.

Additionally, X-band EPR spectroscopy displayed the 7-line hyperfine splitting indicative of a Cu_A site, with *g* values equivalent to the WT protein (Figure 5.9).[13,14] As expected, modifying a single amide in the backbone to an ester does not significantly alter the geometry of the site. However, esters are structurally more flexible than amides because rotation about the C-O-C bonds has a lower barrier and therefore should permit

greater numbers of conformations and potential changes to the engineered site's geometry. That changes in both *g* value and hyperfine splittings are not present suggests that rotation is limited and that the backbone carbonyl interaction therefore likely does not play a significant role in maintaining the site geometry.

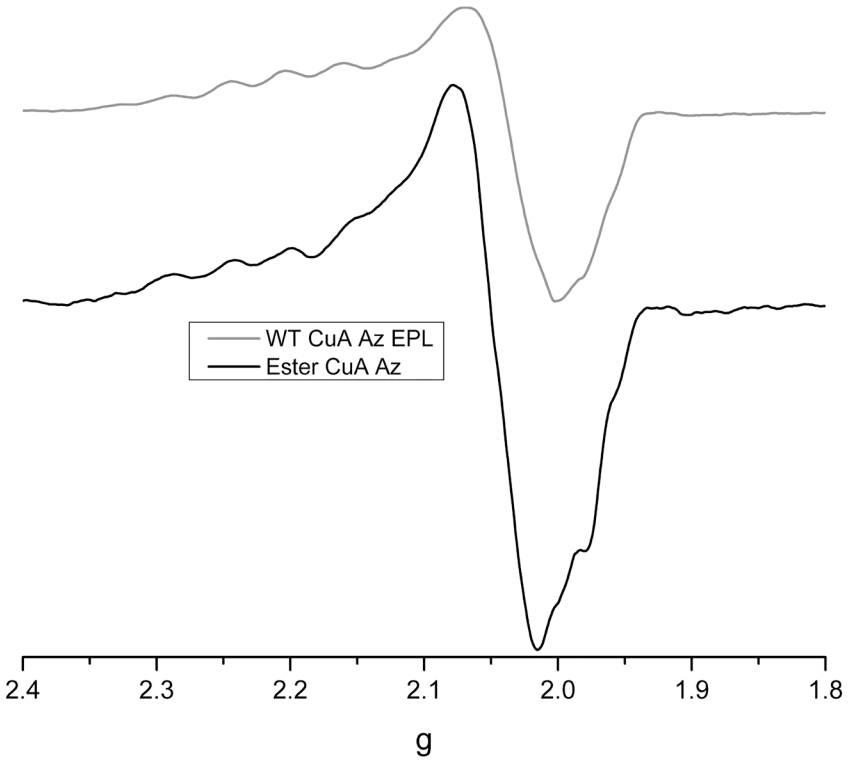


Figure 5.9: X-Band EPR spectrum of Ester Cu_A azurin. Ester Cu_A Az displays the same 7-line hyperfine splitting though not quite as clear as the WT Cu_A Az protein and identical *g* values.

While the geometry of the site likely remained unchanged, changes in the reduction potential could be expected. Because the ester is less structurally rigid than the amide, even small rotations of the ester that are not permissible in the WT protein could alter the reorganization energy by allowing more conformational states of the backbone. Changes in reorganization energy could be reflected in changes in redox behavior.

Additionally, the ester should participate in Cu binding less efficiently than the

amide carbonyl oxygen. To test the effect that the amide-to-ester replacement has on the redox behavior of Cu_A azurin, ester Cu_A was subjected to cyclic voltammetry (CV) following the outlined procedure for CV of WT Cu_A azurin.[16,44] Ester Cu_A displayed the same redox potential as WT Cu_A azurin within experimental error (Figure 5.10). The experimental WT protein redox potential vs. NHE was 277 ± 3 mV and matches reported values.[44] The redox potential determined for the ester variant was 278 ± 5 mV. The similar potentials suggest that the overall contribution from the ester moiety does not substantially alter the interaction of the Glu114-carbonyl to the Cu ion. It further suggests that the amide backbone carbonyl found in WT Cu_A does not impart the redox behavior of the Cu_A site.

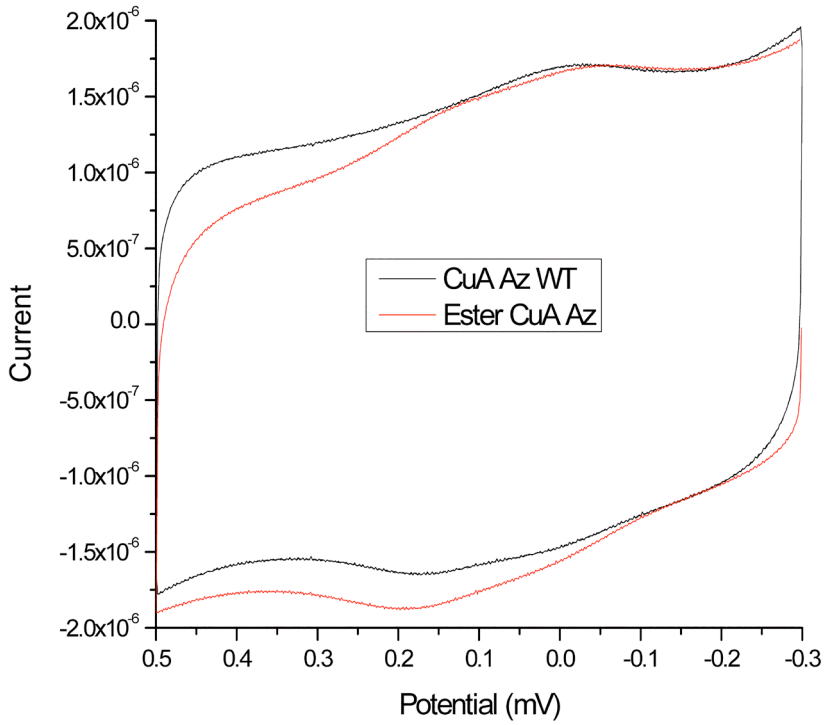


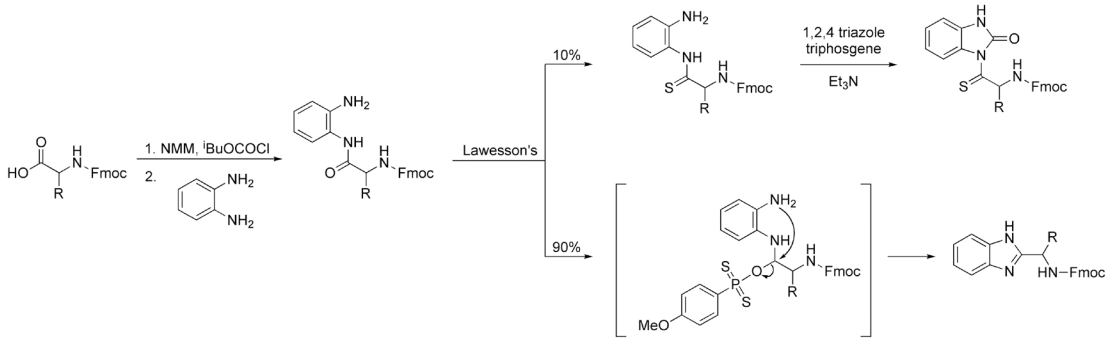
Figure 5.10: CV of Ester Cu_A azurin overlaid with WT Cu_A azurin. The additional peak in the ester spectrum at ~174 mV is attributed to background noise from the electrode and can also be seen in the WT spectrum.

5.8 Installing a Thioamide Into Cu_A Azurin

Inclusion of thioamide moieties into peptides creates peptidomimetics termed thioxopeptides. Similar to depsipeptides, interest in installing thioamide moieties into peptides has been geared towards modifying peptide stability against protease degradation.[45,46] Because thioamides have different hydrogen bonding properties than amides, thioxopeptides are used to investigate H-bonding in α-helices.[31,47,48] The thioamide hydrogen is a stronger hydrogen bond donor than the amide hydrogen, but the sulfur is a weaker hydrogen bond acceptor than the oxygen.[39,48,49] However, in copper coordination chemistry, sulfur is a stronger ligand because it is softer than oxygen and thus a thioamide could more readily and strongly bind copper into the Cu_A site.[50] Though many different routes exist for thioacylation of amino acids (including Lawesson's reagent), many of these routes suffer from lack of specificity, low yield, and racemization.[51-54] A popular methodology that uses phosphorous pentasulfide (P_4S_{10}), specifically thionates amide carbonyls while maintaining high yields.[47]

Incorporating thioamide analogues using SPPS necessitates that the amino acid derivates be synthesized as preformed activated species.[55] Two problems are commonly associated with synthesizing preformed activated species of thioamides, racemization and poor yields.[47] Early synthetic schemes using preformed benzimidazolinones suffer from significant benzimidazole formation and therefore poor overall yields (<20%) (Scheme 5.4).[47,56,57] These syntheses also relied on Lawesson's reagent rather than P_4S_{10} and thus typically resulted in substantial levels of byproduct.[47,56,57] Additionally, these schemes often result in racemized final products and decreased enantiomeric purity. Synthesizing preformed benzotriazoles can circumvent these problems because it is a very good

leaving group.[47,58] Moreover, use of P_4S_{10} reduces byproduct formation found when using Lawesson's reagent. Adding electron-withdrawing groups *para* to the unacylated amino group reduces the amine's nucleophilicity and inhibits its participation in forming the benzimidazole. The benzotriazole also helps reduce racemization (Scheme 5.4).[47]

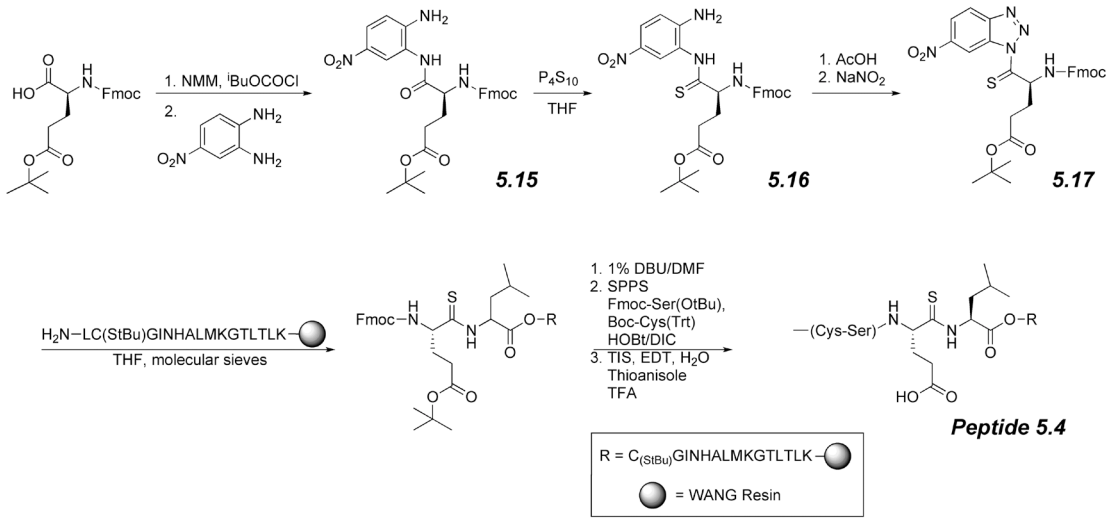


Scheme 5.4: Synthesis of benzimidazolinones for incorporation of thioamides into peptides. Use of Lawesson's reagent results in 90% yield of benzimidazole byproduct and is therefore a poor route to obtain thioamide derivatives.

Preparing a preformed benzotriazole was chosen to synthesize a thioacylated Glu derivative (Glu(S)). Using mixed carbonic anhydride methodology for peptide synthesis, the starting material, Fmoc-Glu(OtBu)-OH, was coupled with 4-nitro-1,2-phenylenediamine to afford anilide ***5.15*** in a near quantitative yield (97%) (Scheme 5.5, Step 1). Thioacylation of ***5.15*** was completed with P_4S_{10} in the presence of Na_2CO_3 to synthesize the thioamide ***5.16*** (Scheme 5.5, Step 2). Though no purification of thioamide-intermediates in the original published scheme was noted,[47] a Glu(S) derivative was not reported and significant byproducts and excess P_4S_{10} were noticeable by ^{1}H NMR spectroscopy and TLC. An additional spot, attributed to acylthioimidate via competitive cyclization of the Glu side chain carbonyl with the nucleophilic thioamide sulfur was also found.[47] The impurities were not further characterized and were removed by silica gel

flash chromatography before proceeding to the final compound as the byproducts proved problematic during later synthetic steps.

The final nitro-benzotriazole product ***5.17*** was formed through intramolecular diazonium cyclization using nitrous acid generated *in situ* with acetic acid and $NaNO_2$ (Scheme 5.5, Step 3). It should be noted that the final nitro-benzotriazole product is not stable for significant periods of time at -20 ˚C and should be used soon after synthesis.



Scheme 5.5: Synthesis of Glu(S) thioamide derivative for installation of a thioamide in Cu_A azurin. See text for details. Modified from Ref. 23.

Standard Fmoc-based SPPS was used to synthesize the first 16 amino acids (H_2N-$LC_{(StBu)}$GINHALMKGTLGLK-OH) of the peptide ***5.4*** (Figure 5.11). Following Fmoc-deprotection of Leu^{115}, compound ***5.17*** was added to the resin bound peptide in dry THF and activated 2 Å molecular sieves and stirred rapidly in a round bottom glass flask for 5 h or until complete coupling was noted by qualitative ninhydrin testing. Once complete, the peptide was quickly deprotected with 20% piperidine and the synthesis finished by coupling Ser113 and Cys112. Following cleavage however, the truncated peptide missing

Cys112 - Glu(S)114 was the primary product. Deprotection of Glu(S)114 and Ser113 with 1% DBU and installation of Boc-amine protected Cys112 was sufficient to eliminate formation of truncated peptide.[37,38]

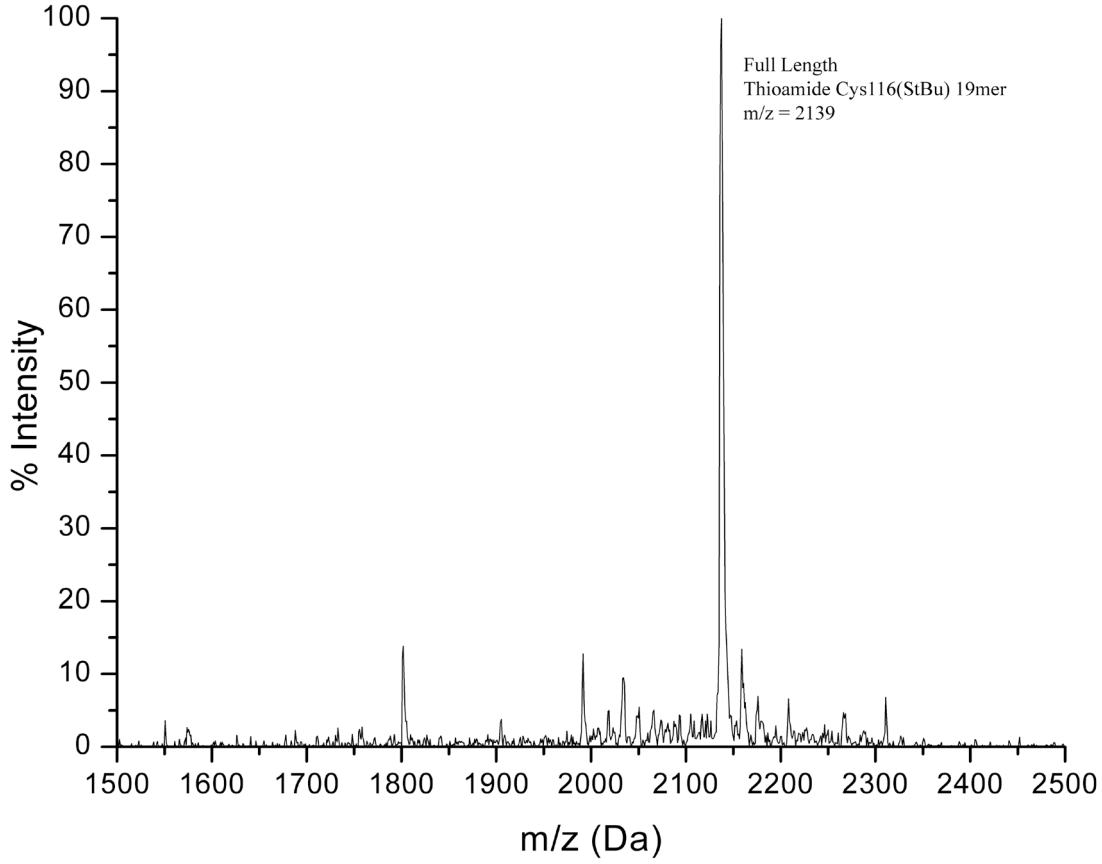


Figure 5.11: MALDI-MS of FL peptide ***5.4*** via Fmoc-SPPS. FL peptide m/z = 2139. Less than 10% of the truncated peptide at Leu[115] is found when using DBU to deprotect the Fmoc from residues after Leu115.

5.9 SEMI-SYNTHESIS OF THIOAMIDE-Cu_A AZURIN BY EPL

The FL peptide derived from Fmoc-SPPS was purified and used in EPL reactions following the same methodology described for the WT Cu_A Az reactions. No pH sensitivity of the thioamide was noticeable. The FL apo protein from EPL reactions was immediately refolded in the presence of DTT after ligation and titrated with $Cu(II)SO_4$.

The protein was titrated with 0.7 equiv. Cu(II)SO_4 when the protein turned pale purple in color. Holo-thioamide Cu_A Azurin showed three LMCT bands though they were weak. The LMCT bands were at 455 nm, 510 nm, and a *d-d* transition at 730 nm (Figure 5.12), and all were red-shifted from the WT bands at 475 nm, 530 nm, and 750 nm. The thioamide protein LMCT peaks displayed much greater shift than those of the ester Cu_A Az variant, suggesting that the additional thiocarbonyl from the thioamide may indeed be binding the Cu stronger.

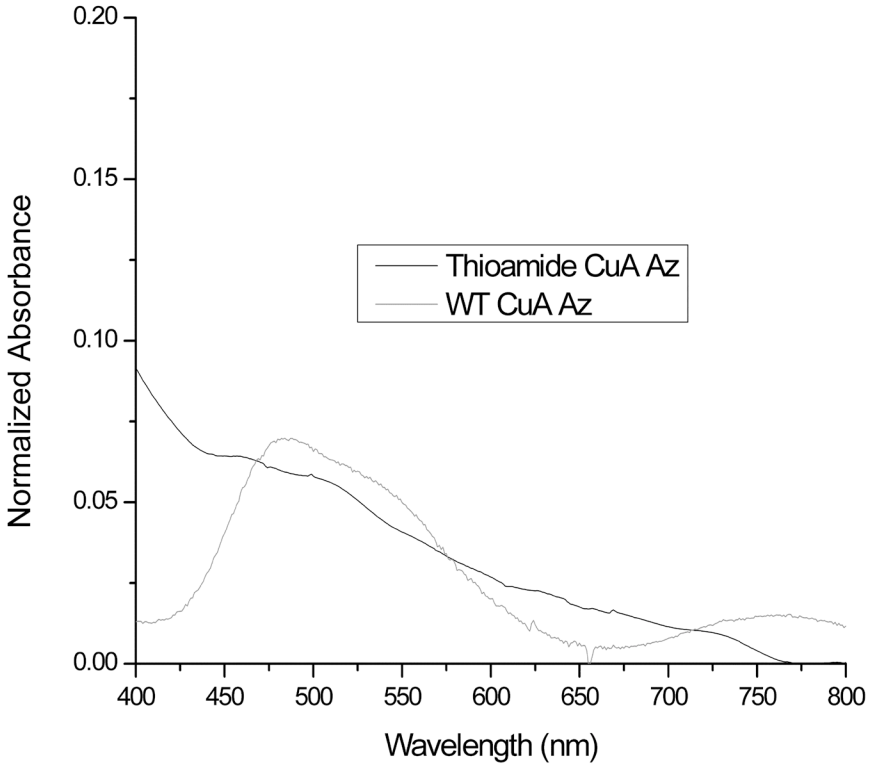


Figure 5.12: UV-Vis overlay of WT EPL Cu_A azurin and Thioamide Cu_A azurin. The LMCT bands are red shifted for the thioamide protein (455 nm, 510 nm, and 730 nm). The WT LMCT bands are at 480 nm, 530 nm and 775 nm.

Unfortunately insufficient protein was obtained through EPL reactions to afford a concentration acceptable for EPR experimentation; something that will need to be examined before coming to major conclusions about this protein.

5.10 CONCLUSIONS

Incorporation of nonproteinogenic amino acid backbone analogues into engineered Cu_A azurin has successfully been completed. Correct incorporation of an ester and thioamide moiety into the scaffold to alter the behavior of the Glu114 backbone carbonyl interaction with Cu was achieved using EPL. The insertions were designed to minimally perturb the geometry of the Cu_A site to specifically isolate any potential electronic interactions the Glu114 carbonyl had with Cu.

Incorporating an ester at Glu114 deleted a hydrogen bond partner and was originally thought to possibly alter geometry because the hydrogen bonding of the residue may be key to correct folding. Additionally, increased rotational flexibility about the ester compared to the native amide was thought to alter the geometry of the site. However, EPR analysis and determination of the variant's redox behavior showed that the geometry of the site was very similar to WT Cu_A azurin with similar *g* values found in the EPR spectrum and a virtually identical reduction potential for the ester variant and WT Cu_A azurin.

Moreover, the ester variant displayed a red-shifted UV-Vis spectrum when compared to the WT Cu_A Az scaffold under identical conditions. One of the LMCT bands found at 480 nm in WT Cu_A Az was red-shifted to 475 nm and the *d-d* transition between the bound Cu ions was red-shifted from 775 nm in the WT scaffold to 750 nm.

These shifts suggest increased energy in binding of the Cu though UV-Vis spectra are not determinative of binding affinity. Regardless, CV did not confirm any potential increased binding affinity since no redox changes were observed. While new data has established that the Glu114 carbonyl does not appear to play a significant role in instituting correct geometry or in determining the redox behavior of the Cu_A site in engineered Cu_A Az, it appears that on at least some level, the amide-to-ester replacement affects the binding of Cu from UV-Vis studies.

A synthetic scheme for successful incorporation of thioamides into Cu_A Az has been devised. This scheme is capable of making thioamide derivatives in a minimal number of steps with high overall yields of pure product. However, the final activated synthesized species is unstable and degrades over extended time. Regardless, successful peptide synthesis and ultimate semi-synthesis of a Cu_A Az variant containing an amide-to-thioamide replacement has been completed. Insertion of a thioamide was expected to cause significant disruption of the Cu_A features because the S is much softer than O and a preferred ligand in Cu coordination.[50] Though extremely preliminary at this time, the UV-Vis spectroscopy displays dramatic shifts of LMCT bands that could be associated with distorted Cu binding and altered Cu binding affinity.

5.11 Future Directions

Before coming to major conclusions about the effect the amide-to-ester replacement has on the Cu_A scaffold, more detailed information regarding the structure is necessary. While crystallization of EPL proteins is notoriously difficult due to low yields of pure holo-protein when compared to yields available through recombinant protein

expression, other spectroscopic methods such as ENDOR and ESEEM could help establish changes in bond lengths and angles demonstrating geometric changes to the site. Previous collaborations with the Blackburn lab at Oregon have provided such insights in other EPL metalloproteins semi-synthesized in our laboratories.

Disappointingly, yields of the thioamide Cu_A Az protein are low. Because of such low yields, no spectroscopic work other than UV-Vis studies has been completed at this time. The thioamide UV-Vis spectrum shows strongly red-shifted LMCT bands suggesting that the additional S may bind Cu stronger than its natural amide counterpart. However, the data are very preliminary and conclusions cannot be made about the amide-to-thioamide replacement with the lack of other evidence available at this time. Thus what, if any effect the amide-to-thioamide replacement has on Cu binding efficiency, geometry, and redox behavior must be further investigated using a wide range of techniques including EPR, ENDOR, and CV.

5.12 EXPERIMENTAL

General. All chemicals used were obtained from Aldrich. All amino acids were purchased from Novabiochem, Advanced Chemtech, or Chem-Impex International. Protecting groups used for Fmoc-SPPS were Cys(Trt), Ser(OtBu), Glu(OtBu), Cys(StBu), Asn(Trt), His(Trt), Lys(Boc), Thr(OtBu). Protecting groups used for Boc-SPPS were Cys(Trt), Ser(OBzl), Glu(OBzl), Cys(StBu) Asn(Xan), His(DNP), Lys(Z), Thr(OBzl). Methylene chloride (DCM) was distilled from calcium hydride. Tetrahydrofuran (THF) was distilled from sodium and benzophenone. All aqueous buffers were degassed under Ar or N_2 as specified immediately prior to use.

All NMR spectra were obtained with a Varian U500 NMR Spectrometer in $CDCl_3$. UV-Vis spectra were taken on an HP Diode Array Spectrometer or a Cary 5000 spectrometer. X-band EPR spectra were collected on a Varian E-122 spectrometer at the Illinois EPR Research Center (IERC). The samples were run at ~30 K using liquid He and an Air Products Helitran cryostat without glassing agents. The magnetic field was calibrated using a Metrolab PT 2025 NMR Gaussmeter, and the microwave frequency was measured with an EIP model 578-frequency counter equipped with a high-frequency option.

Mass spectral data was collected by the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois by either Electrospray Ionization (ESI MS), or matrix assisted laser desorption ionization (MALDI) techniques as indicated. Mass spectral data are reported as – mass, peak (percent).

Preparative RP-HPLC was performed with a Waters Delta 600 system. Solution A was 0.1 % TFA in H_2O, and solution B was 80 % MeCN/20 % H_2O with 0.1 % TFA. Unless otherwise stated, a linear gradient of 20% to 70% B over 23 min was used for all runs. Solid phase peptide synthesis (SPPS) was performed on a Rainin PS3 automated peptide synthesizer or manually using medium fritted funnels.[59]

H_2N-CSELCGINHALMKGTLTLK-COOH

Synthesis of Peptide *5.1*, Lab Book VI, page 24.

The peptide was synthesized on a 0.1 mmol scale. Fmoc Wang resin preloaded with Fmoc-Lys(Boc) was pre-swelled in DMF (6 x 10 min, 6 mL). Synthesis of the peptide from Lys130 through Gly117 was completed using a six-fold excess of amino

acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (6 equiv) for 2 min prior to coupling to the resin. All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-Cys(Trt)-OH (4 equiv.) was coupled after Gly117 in the presence of hydroxybenzotriazole (HOBt) (8 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (8 equiv.) in DMF at room temperature (RT) for 2-5 h. After Fmoc-deprotection with 20% piperidine, Leu115, Glu114, and Ser113 were coupled using 6 equiv HBTU in 0.4 M NMM after a 2 min pre-activation. The N-terminal Cys112 was coupled as Boc-Cys(Trt)-OH (4 equiv.) in the presence of HOBt/DIC (8 equiv. each) with N_2 sparging for 5 h. Amino acids were double coupled when necessary and couplings were monitored by qualitative ninhydrin test. Following addition of the last residue, the resin was washed with DCM and dried for 10 min under reduced pressure before the peptide was cleaved from the resin in a cocktail of 2% H_2O (100 μL), triisopropylsilane (TIS) (100 μL), and ethanedithiol (EDT) (100 μL) in neat trifluoroacetic acid (TFA) (5 mL) with stirring at RT for 2 h. The solvent was evaporated and the peptide was precipitated using cold Et_2O to obtain a white solid. The solid was dissolved in a minimal amount of 1:1 H_2O/MeCN (5-10 mL) and lyophilized prior to purification by preparative RP-HPLC using a C18 column to obtain the final products as a white solid after lyophilization. (Yield: 82 mg). MALDI MS. Expected: 2032.45, Observed: 2032.8.

H_2N-CSELC$_{(StBu)}$GINHALMKGTLTLK-COOH

Synthesis of Peptide *5.2*, Lab Book VI, page 137.

The peptide was synthesized on a 0.15 mmol scale. Fmoc Wang resin preloaded with Fmoc-Lys(Boc) was pre-swelled in DMF (6 x 10 min, 6 mL). Synthesis of the peptide from Lys130 through Gly117 was completed using a six-fold excess of amino acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (6 equiv) for 2 min prior to coupling to the resin. All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-Cys(StBu)-OH (4 equiv.) was coupled after Gly117 in the presence of hydroxybenzotriazole (HOBt) (8 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (8 equiv.) in DMF at RT for 2-5 h. After Fmoc-deprotection with 20% piperidine, Leu115, Glu114, and Ser113 were coupled using 6 equiv HBTU in 0.4 M NMM after a 2 min pre-activation. The N-terminal Cys112 was coupled as Boc-Cys(Trt)-OH (4 equiv.) in the presence of HOBt/DIC (8 equiv. each) with N_2 sparging for 5 h. Amino acids were double coupled when necessary and couplings were monitored by qualitative ninhydrin test. Following addition of the last residue, the resin was washed with DCM and dried for 10 min under reduced pressure before the peptide was cleaved from the resin in a cocktail of 2% H_2O (100 μL), triisopropylsilane (TIS) (100 μL), and ethanedithiol (EDT) (100 μL) in neat trifluoroacetic acid (TFA) (5 mL) with stirring at RT for 2 h. The solvent was evaporated and the peptide was precipitated using cold Et_2O to obtain a white solid. The solid was dissolved in a minimal amount of 1:1 H_2O/MeCN and lyophilized prior to purification by

preparative RP-HPLC using a C18 column to obtain the final products as a white solid after lyophilization. MALDI MS. (Yield: 48 mg). Expected: 2120.62, Observed: 2120.3.

H_2N-CSE[ψCOO]$LC_{(StBu)}$GINHALMKGTLTLK-COOH

Synthesis of Peptide *5.3* on Wang Resin, Lab Book VI, page 125.

The peptide was synthesized on a 0.35 mmol scale. Wang resin preloaded with Fmoc-Lys(Boc) was swelled in DMF (6 x 10 min, 6 mL). Synthesis of the peptide from Lys130 through Gly117 was completed using a six-fold excess of amino acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (6.6 equiv) for 2 min prior to coupling to the resin. All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-Cys(StBu)-OH (4 equiv.) was coupled after Gly117 in the presence of hydroxybenzotriazole (HOBt) (8 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (8 equiv.) in DMF at RT for 2-5 h. After Fmoc-deprotection with 20% piperidine, leucic acid (11 equiv) was coupled to the peptide in the presence of HOBt (12.5 equiv.), DIC (11 equiv.), and NMM (4.4 equiv.) at RT for 3 h. Fmoc-Glu(OtBu)-OH (11 equiv.) was triple coupled to the resin in the presence of DIC (11 equiv.), NMM (4.4 equiv.), and 5 mol% 4-dimethylaminopyridine (DMAP) (0.55 equiv.). The completion of the couple was monitored using a hydroxyl specific colorimetric test similar to the Kaiser test (see below). Glu114 and Ser113 were Fmoc-deprotected with 1% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in DMF (w/v). The N-terminal Cys112 was coupled as Boc-Cys(Trt)-OH (4 equiv.) in the presence of HOBt/DIC (8 equiv. each) with N_2 sparging for 5 h. Amino acids were double coupled

when necessary and couplins were monitored by qualitative ninhydrin test. The peptide was cleaved from the resin in a cocktail of 2% triisopropylsilane (TIS) (200 μL), 2% ethanedithiol (200 μL) in neat TFA (10 mL) with stirring at -5 °C for 2 h and at RT for an addition 0.5-1 h. The solvent was evaporated and the peptide precipitated using cold Et_2O to obtain a white solid. The solid was dissolved in a minimal amount of 1:1 H_2O/MeCN (5-10 mL) and lyophilized prior to purification by preparative RP-HPLC using a C18 column to obtain a white solid after lyophilization. (Yield: 18 mg) MALDI MS. Expected: 2121.62, Observed: 2121.45.

H_2N-CSE[ψCOO]$LC_{(StBu)}$GINHALMKGTLTLK-COOH

Synthesis of Peptide *5.3* Using PAM Resin, Lab Book VI, page 187.

The peptide was synthesized on a 1.0 mmol scale. 4-Hydroxymethyl-phenylacetamidomethyl (PAM) resin preloaded with H-Lys(2-Cl-Z) was swelled in DMF (6 x 10 min, 6 mL). Synthesis of the peptide from Lys130 through Gly117 was completed using a four-fold excess of amino acids pre-activated with 20% diisopropylethylamine (DIPEA) in DMF (v/v) for 5 min prior to coupling to the resin. All Boc-deprotections were accomplished using neat trifluoroacetic acid (TFA) with N_2 sparging (1 x 3 min, 6 mL) and the resin was washed with 1:1 DCM/DMF by continual flow through for 1-2 min before completing the next coupling. Boc-Cys(StBu)-OH (4 equiv.) was coupled in the presence of hydroxybenzotriazole (HOBt) (8 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (8 equiv.) in DMF with N_2 sparging for 5 h. Amino acids were double coupled when necessary. Leucic acid (4 equiv.) was coupled in the presence of HOBt (16 equiv), DIC (8 equiv) and NMM (1.6 equiv) at RT for 3 h. Boc-Glu(OBz.)-

OH (7.7 equiv.) was triple coupled to the resin in the presence of DIC (7.7 equiv.), and 5 mol% 4-dimethylaminopyridine (DMAP). The completion of the coupling was monitored using a hydroxyl specific colorimetric test similar to the Kaiser test (see below). The remaining residues (Ser113 and Cys112) were coupled using 4 equiv. amino acid in the presence of HOBt/DIC (8 equiv. each). The N-terminal Boc residue was removed with neat TFA (6 mL) for 10 min prior to cleavage of the peptide from the resin. The resin was then washed iteratively with DCM and DMF and dried for 10 min under reduced pressure. The peptide was cleaved from the resin in a cocktail of 10% trifluoromethanesulfonic acid (TFMSA) (400 μL), 2.5% dH_2O (100 μL), triisopropylsilane (100 μL), and thioanisole (100 μL) in neat TFA (4 mL) with stirring at RT for 3 h. The solvent was evaporated carefully under reduced pressure and the resultant oil redissolved in 2:1 THF/TFA (15 mL). The organic layer was then extracted with dH_2O (3 x 15 mL) and the aqueous layers combined and extracted with Et_2O. The aqueous layer was then lyophilized to yield a white solid. The solid was purified by preparative RP-HPLC using a C18 column to obtain a white solid after lyophilization. (Yield: 55 mg) MALDI MS. Expected: 2121.62, Observed: 2121.51.

Hydroxyl specific test for SPPS[40]

The resin with peptide was washed with dry DCM (5 x 25 mL). A few beads of resin were removed and placed into a small flint tube (Sample 1). Two other flint tubes were used as controls. Sample 2 contained fully unprotected alanine in DMF. Sample 3 contained fully unprotected leucic acid in DMF. To each sample was added 3 drops of 0.01 M tosyl chloride (TsCl) in dry toluene followed by addition of 3 drops of 0.025 M 4-

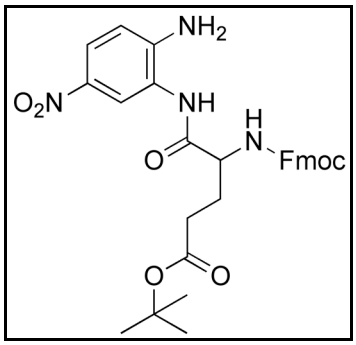
(4-nitrobenzyl) pyridine (PNBP). The orange colored solutions were heated using a handheld heat gun until the solution turned clear. Without cooling, 5 drops of 10% piperidine in $CHCl_3$ (v/v) was added. Positive tests for unprotected hydroxyl groups appear pink/purple in color similar to Sample 3. Negative tests for unprotected hydroxyls appear clear in color similar to Sample 2. Double coupling of leucic acid was completed when samples appeared pink/purple in color.

H_2N-CSE[ψCSN]LC$_{(StBu)}$GINHALMKGTLTLK-COOH

Synthesis of Peptide *5.4*, Lab Book VI, page 127.

The peptide was synthesized on a 0.2 mmol scale. Wang resin preloaded with Fmoc-Lys(Boc) was swelled in DMF (6 x 10 min, 3 mL). Synthesis of the peptide from Lys130 through Gly117 was completed using a six-fold excess of amino acids pre-activated with 0.4 M N-methylmorpholine (NMM) (6 mL) and 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) (6 equiv.) for 2 min prior to coupling to the resin. All Fmoc-deprotections were accomplished using 20% piperidine/DMF (v/v) (3 x 3 min, 6 mL). Fmoc-Cys(StBu)-OH (4 equiv.) was coupled after Gly117 in the presence of hydroxybenzotriazole (HOBt) (8 equiv.) and N,N'-diisopropylcarbodiimide (DIC) (8 equiv.) in DMF at RT for 2-5 h. Fmoc-Leu-OH (6 equiv.) was coupled in the presence of HOBt/DIC (8 equiv. each) in DMF for 3 h. Amino acids were double coupled when necessary. 1-(N-Fmoc-L-thionoglutyamyl), (compound ***5.17***, 4 equiv.) was dissolved in dry THF (4 mL), added to the resin in dry THF (4 mL) with stirring, and stirred under Ar for 4 h. The resin was then washed with fresh dry THF and the amino acid coupled a second time for 4 h. Coupling was

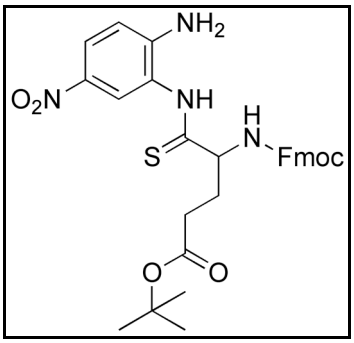
monitored by qualitative ninhydrin test. All Fmoc-deprotections after installation of the thioamide were completed with 1% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in DMF (w/v). Boc-Cys(Trt)-OH (4 equiv.) was coupled in the presence of HOBt/DIC (8 equiv. each). The resin was washed with DCM and dried for 10 min under reduced pressure. The peptide was cleaved from the resin in a cocktail of 82.5%/5%/5%/5%/2.5% TFA/thioanisole/Triisopropylsilane/dH_2O/ethanedithiol (5 mL total volume) and stirred at RT for 2 h. The solvent was evaporated and the peptide precipitated using cold Et_2O to obtain a white solid. The solid was dissolved in a minimal amount of 1:1 H_2O/MeCN and lyophilized prior to purification by preparative RP-HPLC using a C18 column to obtain a white solid after lyophilization. (Yield: 33 mg) MALDI MS. Expected: 2136.69, Observed: 2137.1.



Fmoc-Glu(OtBu)-2-amino-5-nitroanilide (***5.15***)**,** Lab Book VI , page 118.

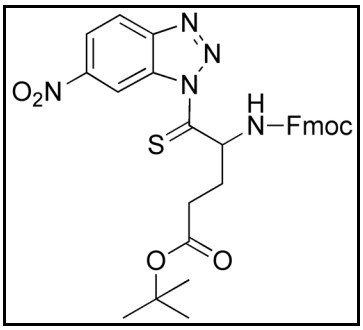
To a solution of Fmoc-Glu(OtBu)-OH (2.5 g, 5.88 mmol) dissolved in dry THF (20 mL) and cooled to -20 °C was added N-methylmorpholine (NMM) (1.2 g, 11.8 mmol) followed by the dropwise addition of 1 equiv. of ethyl chloroformate (638 mg, 5.88 mmol). The reaction was stirred at -20 °C under nitrogen for 10 min. A separate flask charged with 1 equiv. of 4-nitro-1,2-phenylenediamine (900 mg, 5.88 mmol) was

purged with Ar and dry THF (20 mL) was added. The flask was cooled to -20 °C and half of the flask contents were transferred to the flask containing the amino acid by cannulation under Ar pressure. The reaction was then stirred at -20 °C for 2 min before transferring the contents of the the amino acid flask to the remaining 4-nitro-1,2-phenylenediamine by cannulation under Ar pressure. The reaction was then stirred at -15 °C for 2 h. The flask was removed from the bath and warmed to RT overnight (16 h). The solvent was removed by rotary evaporation and the resulting oil dissolved in EtOAc. The crude product was washed with sat. NaH_2PO_4 (aq), 5% $NaHCO_3$, and brine. The product was dried under reduced pressure to afford the final product as a yellow solid without further purification. (Yield: 3.11 g, 95%) ^{1}H NMR (500 MHz, $CDCl_3$) δ 1.48 (s, 9 H), 2.18 (d, 2 H, *J* = 48 Hz) 2.58 (d, 2 H, *J* = 67.5 Hz), 4.2 (t, 1 H, *J* = 7 Hz), 4.25 (br s, 1 H), 4.43 (d, 2 H, *J* = 6 Hz), 6.13 (br s, 1 H), 6.68 (d, 1 H, *J* = 10 Hz), 7.36-7.26 (m, 2 H), 7.38 (t, 2 H, *J* = 5.5 Hz), 7.58 (d, 2 H, *J* = 10 Hz), 7.76 (d, 2 H, *J* = 5 Hz), 7.98 (d, 1 H *J* = 15 Hz), 8.17 (d, 1 H, *J* = 57.5 Hz). ^{13}C NMR (125.6 MHz, $CDCl_3$) δ 28.2 (CH_3), 30.5 (CH_2), 32.1 (CH_2), 47.2 (CH_2), 55.6 (CH), 64.8 (CH), 67.5 (CH), 82.1 (Cq), 115.4 (Cq), 120.2 (CH), 123.2 (CH), 124.2 (CH), 125.1 (CH), 127.3 (CH), 128.1 (CH), 138.7 (CH), 141.5 (CH), 143.7 (Cq), 149.1 (Cq), 158.3 (Cq), 170.8 (Cq).



Fmoc-Glu(OtBu)-2-amino-5-nitrothioanilide (***5.16***), Lab Book VI, page 122.

A round bottom flask was charged with 0.5 equiv. of P_4S_{10} (793 mg, 1.8 mmol) and 0.5 equiv. of Na_2CO_3 (189 mg, 1.8 mmol) and purged with Ar. Dry THF (37 mL) was added and the flask contents were stirred at RT for 1 h. The flask was cooled to 0 °C before adding compound ***5.15*** dissolved in dry THF (20 mL). The reaction was stirred at 0 °C for 30 min and warmed to RT. The reaction was monitored until determined complete by TLC (2 h). The solvent was removed by rotary evaporation and the resulting oil was purified directly by flash chromatography (2:1 EtOAc/Hexane) to obtain the final product as a yellow powder. (Yield: 1.43 g, 70%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 1.48 (s, 9 H), 2.27 (d, 2 H, *J* = 40 Hz), 2.58 (d, 2 H, *J* = 55 H), 4.2 (t, 1 H, *J* = 7 Hz), 4.42-4.36 (m, 2 H), 4.6 (br s, 1 H), 6.32 (br s, 1 H), 6.74 (d, 1 H, *J* = 10 Hz), 7.32-7.28 (m, 2 H), 7.38 (t, 2 H, *J* = 7.5 Hz), 7.58 (d, 2 H, *J* = 7 Hz), 7.77 (d, 2 H, *J* = 7 Hz), 8.07 (d, 1 H, *J* = 7 Hz), 8.13 (s, 1 H). ^{13}C NMR (125.6 MHz, $CDCl_3$) δ 28.2 (CH_3), 30.3 (CH_2), 32.1 (CH_2), 47.1 (CH_2), 61.9 (CH), 66.1 (CH), 67.7 (CH), 82.1 (Cq), 115.4 (Cq), 120.31 (CH), 125.1 (CH), 125.7 (CH), 127.3 (CH), 128.1 (CH), 138.6 (CH), 141.1 (CH), 143.6 (CH), 148.5 (Cq), 157.1 (Cq), 173.7 (Cq), 205.2 (Cq).



1-(N-Fmoc-L-thionoglutamyl)-6-nitrobenzotriazole (3), Lab Book VI, page 123.

To a solution of compound ***2*** (1 g, 1.7 mmol) in 1:1 THF/AcOH (19 mL) was added 1.5 equiv. of $NaNO_2$ (180 mg, 2.6 mmol). The reaction was stirred at RT for 75 min and quenched with 3:1 EtOAc/Heptane (30 mL). The organics were then washed with dH_2O, sat $NaHCO_3$, and brine and dried over $MgSO_4$. The solvent was removed by rotary evaporation to afford the final compound as a yellow powder which was used without further purification. (Yield: 908 mg, 90%) ^{1}H NMR (500 MHz, $CDCl_3$) δ 1.48 (s, 9 H), 2.27 (d, 2 H, *J* = 40 Hz), 2.58 (d, 2 H, *J* = 55 H), 4.2 (t, 1 H, *J* = 7 Hz), 4.42-4.36 (m, 2 H), 4.6 (br s, 1 H), 6.32 (br s, 1 H), 7.32-7.28 (m, 2 H), 7.38 (t, 2 H, *J* = 7.5 Hz), 7.58 (d, 2 H, *J* = 7 Hz), 7.77 (d, 2 H, *J* = 7 Hz), 8.22 (d, 1 H, *J* = 10 Hz), 8.38 (d, 1 H, *J* = 10 Hz), 8.89 (s, 1 H).

General procedure for expressed protein ligation, Lab Book VI, page 141.

Cultures of *E. coli* BL21 (DE3) cells containing a plasmid to express the azurin(1-111)-Intein-CBD fusion protein were grown in LB media for 8 h at 37 ºC and were used to inoculate eight 2 L flasks of LB media containing 100 mg/mL ampicillin. The cells were grown at 37 ºC for 16 h with shaking at 210 rpm. Protein expression was induced at ~16 h with 0.5 mM IPTG and induction was continued for 4 h at 37 ºC. The cells were then harvested at 4,000 x g, stored at -20 ºC and used when needed.

Frozen azurin-thioester cell stock was re-suspended in a lysis buffer containing 20 mM HEPES, pH 7.2, 250 mM NaCl, 1 mM EDTA, 1 mM PMSF 0.1% Triton-X-100, and 1 M urea. The suspension was then lysed using sonication (Misonix Sonicator 4000, 0.5 inch diameter probe) for a total work time of 9 min (6 on, 10 rest). The lysate was centrifuged at 13,250 x g for 30 min. at 4 °C. The fusion protein was then bound by batch absorption to 100 mL of chitin resin that had been pre-equilibrated with 20 mM HEPES, pH 7.2, 250 mM NaCl, and 1 mM EDTA (buffer 1) for 1-2 h at 4 ºC. The chitin resin was then poured into a column and the column headspace was purged with Ar. The column was then washed with 3 column volumes of degassed buffer 1 under Ar pressure.

Ligation was initiated by the addition of peptide and in 35 mL of buffer 1 containing 50 mM 4-mercaptophenylacetic acid (MPAA) under Ar and transferred directly to the column by Ar pressure. The chitin resin was then re-suspended in the column and the entire column was agitated gently at 4 ºC for 66 h.

After ligation, the column was eluted under Ar pressure and washed with 1 column volume of buffer 1. The eluent was centrifuged at 13,250 x g for 30 min and the supernatant was concentrated ~5x using 10,000 MWCO Centricon concentration spin tubes at 3000 x g and exchanged into 50 mM NH_4OAc, pH 5.1 *via* G25 sephadex PD-10 columns.

General Procedure for Refolding and Titration, Lab Book 6, page 146.

Dithiothreitol (DTT) (0.9 mM), guanidinium hydrochloride (GdmHCl) (4 M) and $ZnCl_2$ (1 mM) were added directly to a concentrated protein sample (~ 1 mL). After full dissolution, the sample was then left to sit at 25 ºC without disruption for 15 min. The

denatured sample was then added dropwise to a 50x excess volume of 50 mM NH_4OAc buffer, pH 5.1 at 4 ºC containing DTT (0.8 mM). The refolded sample was then exchanged into fresh 50 mM NH_4OAc buffer, pH 5.1 at 4 ºC using G25 Sephadex PD-10 columns.

Following concentration of the sample, the protein was titrated with $Cu(II)SO_4$. Samples titrated with subequivalent additions of Cu(II) were titratated until saturation of the LMCT at 480 nm. The titrated samples were then purified separately using a Pharmacia Q HiTrap HP column (5 mL) which had been washed with 10 column volumes of 1 M NaCl in H_2O and then equilibrated with 10 column volumes of 50 mM NH_4OAc, pH 5.1. Following purification, proteins were concentrated and analyzed using MALDI-TOF-MS and ESI-MS.

Electrochemical Measurements for Ester Cu_A Az, Lab Book 7, page 106.

The reduction potential for the WT Cu_A was verified prior to taking any measurements for the ester Cu_A variant. All measurements were taken by cyclic voltammetry using a CH Instruments 617A potentiostat equipped with a picoamp booster and a faraday cag. A pyrolytic graphite edge (PGE) electrode was polished and soaked in the protein solution (200 μM). The voltage was cycled from 0.5 V to -0.3 V for 3-5 min. The electrode was removed from the protein solution and transferred into fresh 50 mM NH4OAc buffer, pH 5.1 at 4 °C that was degassed with Ar. The electrode was then cycled from 0.5 V to -0.3 V for 5-10 min or until loss of signal. The electrodes were used only a single time before cleaning by sonication and repeating the process.

The reduction potential was also determined by using a film of didodecyldimethylammonium bromide (DDAB) to facilitate protein binding to the electrode. A 20% solution of DDAB in $CHCl_3$ (5 μL) was applied to the electrode and allowed to dry overnight. The electrode was then soaked in the protein solution and the voltage cycled from 0.5 V to -0.3 V for 5-10 min or until loss of signal. This method gave identical results to the previously described method. All reduction potentials were measured against Ag/AgCl and converted to NHE

5.13 REFERENCES

(1) Berry, S. M.; Gieselman, M. D.; Nilges, M. J.; van der Donk, W. A.; Lu, Y. "An Engineered Azurin Variant Containing a Selenocysteine Copper Ligand." *J. Am. Chem. Soc.,* **2002**, *124*, 2084-2085.

(2) Berry, S. M.; Ralle, M.; Low, D. W.; Blackburn, N. J.; Lu, Y. "Probing the Role of Axial Methionine in the Blue Copper Center of Azurin with Unnatural Amino Acids." *J. Am. Chem. Soc.,* **2003**, *125*, 8760-8768.

(3) Garner, D. K.; Vaughan, M. D.; Hwang, H. J.; Savelieff, M. G.; Berry, S. M.; Honek, J. F.; Lu, Y. "Reduction Potential Tuning of the Blue Copper Center in *Pseudomonas aeruginosa* Azurin by the Axial Methionine as Probed by Unnatural Amino Acids." *J. Am. Chem. Soc.,* **2006**, *128*, 15608-15617.

(4) Lu, W.; Qasim, M. A.; Laskowski, M.; Kent, S. B. H. "Probing Intermolecular Main Chain Hydrogen Bonding in Serine Proteinase, Protein Inhibitor Complexes: Chemical Synthesis of Backbone-Engineered Turkey Ovomucoid Third Domain." *Biochemistry,* **1997**, *36*, 673-679.

(5) Lu, W.; Randal, M.; Kossiakoff, A.; Kent, S. B. H. "Probing intermolecular backbone H-bonding in serine proteinase-protein inhibitor complexes." *Chem. Biol.,* **1999**, *6*, 419-427.

(6) Ralle, M.; Berry, S. M.; Nilges, M. J.; Gieselman, M. D.; van der Donk, W. A.; Lu, Y.; Blackburn, N. J. "The Selenocysteine-Substituted Blue Copper Center:

Spectroscopic Investigations of Cys112SeCys *Pseudomonas aeruginosa* Azurin." *J. Am. Chem. Soc.,* **2004**, *126*, 7244-7256.

(7) Olsson, M. H. M.; Ryde, U. "Geometry, Reduction Potential, and Reorganization Energy of the Binuclear Cu_A Site, Studied by Density Functional Theory." *J. Am. Chem. Soc.,* **2001**, *123*, 7866-7876.

(8) DeBeer George, S.; Metz, M.; Szilagyi, R. K.; Wang, H.; Cramer, S. P.; Lu, Y.; Tolman, W. B.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. "A Quantitative Description of the Ground-State Wave Function of CuA by X-ray Absorption Spectroscopy; Comparison to Plastocyanin and Relevance to Electron Transfer." *J. Am. Chem. Soc.,* **2001**, *123*, 5757-5767.

(9) Farver, O.; Lu, Y.; Ang, M. C.; Pecht, I. "Enhanced rate of intramolecular electron transfer in an engineered purple CuA azurin." *Proc. Natl. Acad. Sci. U.S.A.,* **1999**, *96*, 899-902.

(10) Robinson, H.; Ang, M. C.; Gao, Y.-G.; Hay, M. T.; Lu, Y.; Wang, A. H. J. "Structural Basis of Electron Transfer Modulation in the Purple Cu_A Center." *Biochemistry,* **1999**, *38*, 5677-5683.

(11) Solomon, E. I.; LaCroix, L. B.; Randall, D. W.; Gamelin, D. R. "Electronic Structure Contributions to Electron Transfer in Blue Cu and CuA." *J. Biol. Inorg. Chem.,* **2000**, *5*, 16-19.

(12) Sun, D.; Wang, X.; Davidson, V. L. "Redox properties of an engineered purple CuA azurin." *Arch. Biochem. Biophys.,* **2002**, *404*, 158-162.

(13) Hay, M.; Richards, J. H.; Lu, Y. "Construction and characterization of an azurin analog for the purple copper site in cytochrome c oxidase." *Proc. Natl. Acad. Sci. U.S.A.,* **1996**, *93*, 461-464.

(14) Hay, M. T.; Ang, M. C.; Gamelin, D. R.; Solomon, E. I.; Antholine, W. E.; Ralle, M.; Blackburn, N. J.; Massey, P. D.; Wang, X.; Kwon, A. H.; Lu, Y. "Spectroscopic Characterization of an Engineered Purple Cu_A Center in Azurin." *Inorg. Chem.,* **1998**, *37*, 191-198.

(15) Kelly, M.; Lappalainen, P.; Talbo, G.; Haltia, T.; van der Oost, J.; Saraste, M. "Two cysteines, two histidines, and one methionine are ligands of a binuclear purple copper center." *J. Biol. Chem.,* **1993**, *268*, 16781-16787.

(16) Hwang, H. J.; Berry, S. M.; Nilges, M. J.; Lu, Y. "Axial Methionine Has Much Less Influence on Reduction Potentials in a Cu_A Center than in a Blue Copper Center." *J. Am. Chem. Soc.,* **2005**, *127*, 7274-7275.

(17) Xie, X.; Gorelsky, S. I.; Sarangi, R.; Garner, D. K.; Hwang, H. J.; Hodgson, K. O.; Hedman, B.; Lu, Y.; Solomon, E. I. "Perturbations to the Geometric and Electronic Structure of the CuA Site: Factors that Influence Delocalization and Their Contributions to Electron Transfer." *J. Am. Chem. Soc,* **2008**, *130*, 5194-5205.

(18) Solomon, E. I.; Randall, D. W.; Glaser, T. "Electronic structures of active sites in electron transfer metalloproteins: contributions to reactivity." *Coord. Chem. Rev.,* **2000**, *200-202*, 595-632.

(19) Farrar, J. A.; Lappalainen, P.; Zumft, W. G.; Saraste, M.; Thomson, A. J. "Spectroscopic and Mutagenesis Studies on the CuA Centre from The Cytochrome-c Oxidase Complex of *Paracoccus Denitrificans*." *Eur. J. Biochem.,* **1995**, *232*, 294-303.

(20) Hay, M. T.; Lu, Y. "Metal-binding Properties of an Engineered Purple CuA center in azurin." *J. Biol. Inorg. Chem.,* **2000**, *5*, 699-712.

(21) Houser, R. P.; Halfen, J. A.; Young, V. G.; Blackburn, N. J.; Tolman, W. B. "Structural Characterization of the First Example of a Bis(mu-thiolato)dicopper(II) Complex. Relevance to Proposals for the Electron Transfer Sites in Cytochrome c Oxidase and Nitrous Oxide Reductase." *J. Am. Chem. Soc.,* **1995**, *117*, 10745-10746.

(22) So, H. "Anaylsis of Intramolecular Electron Transfer in a Mixed-Valence Cu(I)-Cu(II) Complex Using the PKS Model." *Bull. Korean Chem. Soc.,* **1992**, *13*, 385-387.

(23) Hwang, H. J.; Lu, Y. "pH-dependent transition between delocalized and trapped valence states of a CuA center and its possible role in proton-coupled electron transfer." *Proc. Natl. Acad. Sci. USA,* **2004**, *101*, 12842-12847.

(24) Hwang, H. J.; Nagraj, N.; Lu, Y. "Spectroscopic Characterizations of Bridging Cysteine Ligand Variants of an Engineered $Cu_2(S_{Cys})_2$ Cu_A Azurin." *Inorg. Chem.,* **2006**, *45*, 102-107.

(25) Berry, S. M.; Wang, X.; Lu, Y. "Ligand replacement study at the His120 site of purple CuA azurin." *J. Inorg. Biochem.,* **2000**, *78*, 89-95.

(26) Farver, O.; Hwang, H. J.; Lu, Y.; Pecht, I. "Reorganization Energy of the CuA Center in Purple Azurin: Impact of the Mixed Valence-to-Trapped Valence State Transition." *J. Phys. Chem. B.,* **2007**, *111*, 6690-6694.

(27) Lukoyanov, D.; Berry, S. M.; Lu, Y.; Antholine, W. E.; Scholes, C. P. "Role of the Coordinating Histidine in Altering the Mixed Valency of CuA: An Electron Nuclear Double Resonance-Electron Paramagnetic Resonance Investigation." *Biophys. J.,* **2002**, *82*, 2758-2766.

(28) Gilardi, G. M., G.; Rosato, N.; Canters, G.W.; Finnazzi-Agro, A. "Unique Environment of Trp48 in Pseudomonas aeruginosa Azurin As Probed by Site-Directed Mutagenesis and Dynamic Fluorescence Spectroscopy." *Biochemistry,* **1994**, *33*, 1425-1433.

(29) Coin, I.; Beyermann, M.; Bienert, M. "Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences." *Nat. Protocols,* **2007**, *2*, 3247-3256.

(30) Coin, I.; Dolling, R.; Krause, E.; Bienert, M.; Beyermann, M.; Sferdean, C. D.; Carpino, L. A. "Depsipeptide Methodology for Solid-Phase Peptide Synthesis: Circumventing Side Reactions and Development of an Automated Technique via Depsidipeptide Units." *J. Org. Chem.,* **2006**, *71*, 6171-6177.

(31) Reiner, A.; Wildemann, D.; Fischer, G.; Kiefhaber, T. "Effect of Thioxopeptide Bonds on alpha-Helix Structure and Stability." *J. Am. Chem. Soc.,* **2008**, *130*, 8079-8084.

(32) Hocart, S. J.; Murphy, W. A.; Coy, D. H. "Analogues of growth hormone-releasing factor (1-29) amide containing the reduced peptide bond isostere in the N-terminal region." *J. Med. Chem.,* **1990**, *33*, 1954-1958.

(33) Gallo, E. A.; Gellman, S. H. "Hydrogen-bond-mediated folding in depsipeptide models of beta-turns and alpha-helical turns." *J. Am. Chem. Soc.,* **1993**, *115*, 9774-9788.

(34) Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D.; Oberer, L.; Hommel, U.; Amstutz, R.; Widmer, H. "Probing the Helical Secondary Structure of Short-Chain beta-Peptides." *Helv. Chim. Acta,* **1996**, *79*, 2043-2066.

(35) Yang, J.; Gellman, S. H. "Energetic Superiority of Two-Center Hydrogen Bonding Relative To Three-Center Hydrogen Bonding in a Model System." *J. Am. Chem. Soc.,* **1998**, *120*, 9090-9091.

(36) Kuisle, O.; Quinoa, E.; Riguera, R. "A General Methodology for Automated Solid-Phase Synthesis of Depsides and Depsipeptides. Preparation of a Valinomycin Analogue." *J. Org. Chem.,* **1999**, *64*, 8063-8075.

(37) Galonic, D. P.; van der Donk, W. A.; Gin, D. Y. "Site-Selective Conjugation of Thiols with Aziridine-2-Carboxylic Acid-Containing Peptides." *J. Am. Chem. Soc.,* **2004**, *126*, 12712-12713.

(38) Wade, J. D.; Bedford, J.; Sheppard, R. C.; Tregear, G. W. "DBU as an N alpha-deprotecting reagent for the fluorenylmethoxycarbonyl group in continuous flow solid-phase peptide synthesis." *Pept. Res.,* **1991**, *4*, 194-199.

(39) Lee, H.-J.; Choi, Y.-S.; Lee, K.-B.; Park, J.; Yoon, C.-J. "Hydrogen Bonding Abilities of Thioamide." *J. Phys. Chem. A,* **2002**, *106*, 7010-7017.

(40) Pomonis, G. J.; Severson, R. F.; Freeman, P. J. "A spot test diagnostic of hydroxyl groups." *J. Chromatogr. A.,* **1969**, *40*, 78-84.

(41) Frank, R.; Schutkowsk, M. "Extremely mild reagent for Boc deprotection applicable to the synthesis of peptides with thioamide linkages." *Chem. Comm.,* **1996**, *22*, 2497-2616.

(42) Li, B.; Berliner, M.; Buzon, R.; Chiu, C. K. F.; Colgan, S. T.; Kaneko, T.; Keene, N.; Kissel, W.; Le, T.; Leeman, K. R.; Marquez, B.; Morris, R.; Newell, L.; Wunderwald, S.; Witt, M.; Weaver, J.; Zhang, Z.; Zhang, Z. "Aqueous Phosphoric Acid as a Mild Reagent for Deprotection of tert-Butyl Carbamates, Esters, and Ethers." *J. Org. Chem.,* **2006**, *71*, 9045-9050.

(43) Miel, H.; Rault, S. "Total deprotection of N,N'-bis(tert-butoxycarbonyl)guanidines using SnCl4." *Tetrahedron Lett.,* **1997**, *38*, 7865-7866.

(44) Hwang, H.; Ang, M.; Lu, Y. "Determination of reduction potential of an engineered Cu A azurin by cyclic voltammetry and spectrochemical titrations." *J. Biol. Inorg. Chem.,* **2004**, *9*, 489-494.

(45) Bock, M. G.; DiPardo, R. M.; Williams, P. D.; Pettibone, D. J.; Clineschmidt, B. V.; Ball, R. G.; Veber, D. F.; Freidinger, R. M. "Receptor ligands which

bind the oxytocin receptor with selectivity and high affinity. Chemical modification of a Streptomyces silvensis derived cyclic hexapeptide." *J. Med. Chem.,* **1990**, *33*, 2321-2323.

(46) Brown, D. W.; Campbell, M. M.; Chambers, M. S.; Walker, C. V. "Mono- and dithionopeptide synthesis." *Tetrahedron Lett.,* **1987**, *28*, 2171-2174.

(47) Shalaby, M. A.; Grote, C. W.; Rapoport, H. "Thiopeptide Synthesis. alpha-Amino Thionoacid Derivatives of Nitrobenzotriazole as Thioacylating Agents." *J. Am. Chem. Soc.,* **1996**, *61*, 9045-9048.

(48) Miwa, J. H.; Pallivathucal, L.; Gowda, S.; Lee, K. E. "Conformational Stability of Helical Peptides Containing a Thioamide Linkage." *Org. Lett.,* **2002**, *4*, 4655-4657.

(49) Dudek, E. P.; Dudek, G. O. "Proton magnetic resonance spectra of thiocarboxamides." *J. Org. Chem.,* **1967**, *32*, 823-824.

(50) Belle, C.; Rammal, W.; Pierre, J.-L. "Sulfur ligation in copper enzymes and models." *J. Inorg. Biochem.,* **2005**, *99*, 1929-1936.

(51) Cava, M. P.; Levinson, M. I. "Thionation reactions of lawesson's reagents." *Tetrahedron,* **1985**, *41*, 5061-5087.

(52) Guthrie, D.; Williams, C.; Elmore, D. "Configuration of thionopeptide bond in solution." *Int. J. Pept. Protein Res.,* **1986**, *28*, 208-211.

(53) Ried, W.; Emden, W. V. D. "Aminosaure-thionester und Endothiopeptide, II." *Liebigs Ann. Chem.,* **1961**, *642*, 128-133.

(54) Ried, W.; Schmidt, E. "Reaktionen mit Imidsaureestern, IV. *N*-acylierte alpha-Aminoimidsaureester, Iminodipeptide und Endothiodipeptide." *Liebigs. Ann. Chem.,* **1966**, *695*, 217-225.

(55) Hoeg-Jensen, T.; Havsteen Jakobsen, M.; Olsen, C. E.; Holm, A. "Formation of peptide thioamides by use of Fmoc amino monothioacids and PyBOP." *Tetrahedron Lett.,* **1991**, *32*, 7617-7620.

(56) Zacharie, B.; Martel, R.; Sauve, G.; Belleau, B. "Chemoselective thioacylation of amino acids. Preparation of the four monothiothymopentin analogs and their biological activity." *Bioorg. Med. Chem. Lett.,* **1993**, *3*, 619-624.

(57) Zacharie, B.; SauvÈ, G.; Penney, C. "Thioacylating agents. Use of thiobenzimidazolone derivatives for the preparation of thiotuftsin analogs." *Tetrahedron,* **1993**, *49*, 10489-10500.

(58) Katritzky, A. R.; Rachwal, S.; Hitchings, G. J. "Benzotriazole: A novel synthetic auxiliary." *Tetrahedron,* **1991**, *47*, 2683-2732.

(59) Levengood, M. R.; van der Donk, W. A. "Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies." *Nat. Protocols,* **2007**, *1*, 3001-3010.

AUTHOR'S BIOGRAPHY

Kevin Michael Clark graduated from Worcester Polytechnic Institute in 2004 with a Bachelor of Science degree in Biotechnology after attending Clemson University for two years. Kevin worked as an enzymologist at Abbott Laboratories in Worcester before relocating to Urbana, Illinois, to pursue graduate education in biochemistry. After completing his Ph.D., Kevin will continue pursuing his Juris Doctorate, specializing in intellectual property, at Franklin Pierce Law Center in Concord, New Hampshire, where he is expected to graduate in May 2012.